Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System

Daewon Byun[1] and Kenneth L. Schere[2*]

Atmospheric Sciences Modeling Division, Air Resources Laboratory,
National Oceanic and Atmospheric Administration,
Research Triangle Park, NC 27711

[1]Present affiliation: University of Houston, 4800 Calhoun, Houston, TX 77204-5007
[2]On assignment to the National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC

Accepted by Applied Mechanics Reviews

September 2004

ABSTRACT

The Community Multi-scale Air Quality (CMAQ) modeling system has been designed to approach air quality as a whole by including state-of-the-science capabilities for modeling multiple air quality issues, including tropospheric ozone, fine particles, acid deposition, and visibility degradation. CMAQ was also designed to have multi-scale capabilities so that separate models were not needed for urban and regional scale air quality modeling. By making CMAQ a modeling system that addresses multiple pollutants and different spatial scales, CMAQ has a "one atmosphere" perspective that combines the efforts of the scientific community. To implement multi-scale capabilities in CMAQ, several issues, such as scalable atmospheric dynamics and generalized coordinates, that depend on the desired model resolution are addressed. A set of governing equations for compressible non-hydrostatic atmospheres is available to better resolve atmospheric dynamics at smaller scales. Because CMAQ is designed to handle scale-dependent meteorological formulations and a large amount of flexibility, CMAQ's governing equations are expressed in a generalized coordinate system. This approach ensures consistency between CMAQ and the meteorological modeling system. The generalized coordinate system determines the necessary grid and coordinate transformations, and it can accommodate various vertical coordinates and map projections.

The CMAQ modeling system simulates various chemical and physical processes that are thought to be important for understanding atmospheric trace gas transformations and distributions. The CMAQ modeling system contains three types of modeling components (Models-3): a meteorological modeling system for the description of atmospheric states and motions, emission models for man-made and natural emissions that are injected into the atmosphere, and a chemistry-transport modeling system for simulation of the chemical transformation and fate. The CMAQ chemical transport model includes the following process modules: horizontal advection, vertical advection, mass conservation adjustments for advection processes, horizontal diffusion, vertical diffusion, gas-phase chemical reactions and solvers, photolytic rate computation, aqueous-phase reactions and cloud mixing, aerosol dynamics, size distributions and chemistry, plume chemistry effects, and gas and aerosol deposition velocity estimation. This paper describes the Models-3 CMAQ system, its governing equations

LIST OF SYMBOLS

A_j	area of vegetation type *j*
C_{vd}	specific heat capacity for dry air at constant volume,
E_f	emissions factor
E_i	emission rate for species *i*
$F(\lambda)$	actinic flux
F_{ij}	environmental factor
$F_{q_i}^x$, $F_{q_i}^y$, $F_{q_i}^z$	turbulent flux components in Cartesian coordinates
$F_{q_i}^x\big\|_{dep}$, $F_{q_i}^y\big\|_{dep}$	side-wall deposition flux terms
$\hat{F}_{q_i}^k$	turbulent flux component in a generalized coordinate
$\hat{\mathbf{F}}_s$	horizontal forcing vector
$\hat{\mathbf{F}}_{q_i}$	horizontal turbulent flux vector of trace species
f	Coriolis factor
f_c	cloud fraction
f_{sc}	scattering phase function asymmetry factor
g	gravity
H_i	Henry's Law coefficient for species *i*
h	height
h_{AGL}	height above the ground level
h_{dep}	thickness of lowest model layer
i	trace species index
J_{above}	above-cloud photolysis rate
J_{below}	below-cloud photolysis rate
J_{clear}	clear-sky photolysis rate
J_i	photolysis rate
J_s	vertical factor of the Jacobian, $J_s = \lvert\partial h / \partial s\rvert$
j	space index or species index
K_H	horizontal eddy diffusivity in Cartesian coordinate
K_{Hf}	a uniform horizontal eddy diffusivity at a fixed resolution
K_{HT}	horizontal eddy diffusivity representing transport effects
K_{HN}	horizontal eddy diffusivity compensating numerical diffusion
K_{zz}	vertical eddy diffusivity
$\hat{K}^{11}$, $\hat{K}^{22}$, $\hat{K}^{33}$	components of the contravariant eddy diffusivity
$\hat{K}_H^{\max}$	domain maximun contravariant eddy diffusivity

k_l	reaction constant for the *l*-th reaction equation
L	liquid water content
L_i	loss ratio of species *i* due to chemical reactions
M_k	*k*-th moment of particle distribution
M_0	*0*-th moment (total number) of particles
M_2	*2*-nd moment (total particle) surface area
M_3	*3*-rd moment (total particle) volume
m	map scale factor
m_i	molar mixing ratio for species *i*
m_{ibg}	background concentration for species *i*
m_i^{cld}	in-cloud concentration for species *i*
m_{oi}	initial molar mixing ratio for species *i*
m_p	concentration of sub-grid plume
$\overline{m}_i$	mean molar mixing ratio for species *i*
N	total number of particles
n	particle number density
P_r	precipitation rate
p	atmospheric pressure
p_{oo}	reference pressure (10^5 Pascal)
Q_{aero}	external source/sink of aerosol
Q_ρ	mass consistency error term
Q_ζ	source of entropy density
Q_{φ_i}	source of pollutant mass
$\hat{Q}_{m_i}$	source term in molar mixing ratio
q	mixing ratio
$\overline{q}_i$	mean mixing ratio of species *i*
q_i'	turbulent component of species *i* mixing ratio
R_{aero}	particle formation, growth, and depletion rate
R	universal gas constant
R_d	dry air gas constant
R_{φ_i}	chemistry reactions represented in mixing ratio
$\hat{R}_{m_i}$	gas reactions represented in molar mixing ratio
r_l	rate of reaction
S	solar radiation
S_Γ	strength of stretching wind pattern
S_Λ	strength of shearing wind pattern
s	generalized meteorological vertical coordinate
T	air temperature
T_{oo}	reference temperature
t	time in Cartesian coordinate space
t_r	cloud transmissivity
$\hat{t}$	time in generalized coordinate space
u	wind component in x-direction

$\hat{\mathbf{v}}_s$	horizontal wind vector in the generalized coordinate
$\mathbf{v}_z$	horizontal wind vector in the Cartesian coordinate
v	wind component in y-direction
v_{dep}	dry deposition velocity
$\hat{v}^j$	contravariant wind component
$\hat{v}_g$	covariant sedimentation velocity
$\hat{v}_k$	covariant wind component
$\overline{\hat{v}^k}$	mean component of contravariant wind
$(\hat{v}^k)'$	turbulent component of contravariant wind
W	liquid water path
$\overline{W}_T$	mean total water content
w	vertical wind component
$x,$	Cartesian coordinate in east direction
$\hat{x}^1,\ \hat{x}^2,\ \hat{x}^3$	generalized coordinates
y	Cartesian coordinate in north direction
$z,$	Cartesian coordinate in vertical direction
z_{sfc}	topographic height

α_i	scavenging coefficient for species i
α_o	0.28 (used in Smargorinsky horizontal diffusion algorithm)
β_{hdiff}	Courant number for horizontal diffusion
Δt_{sync}	synchronization time step
Δx	horizontal grid size
Δx_f	fixed size of horizontal grid
Δz_{cld}	thickness of cloud
δ_{cld}	cloud optical depth
δV	grid cell volume
δV_p	volume of plume in a grid cell
Φ	geopotential height
ϕ_i	quantum yield
$\hat{\gamma}$	determinant of metric tensor
$\hat{\gamma}_{jk}$	metric tensor
η	eta coordinate
φ_i	concentration of trace species i
φ_i^*	scaled (coupled) trace species concentration
φ_i'	turbulent component of trace species concentration
$\overline{\varphi}_i$	mean concentration of trace species i
λ	wavelength
θ	zenith angle

ρ	air density
ρ'	turbulent component of air density
ρ_{H2O}	density of water vapor
ρ_{oo}	air density of reference atmosphere
$\overline{\rho}$	mean air density
σ_i	absorption cross section
$\sigma_z, \sigma_{\tilde{z}}, \sigma_{\tilde{p}}, \sigma_{p_o}$	terrain-following vertical coordinates
τ_{ulim}	upper time step limit in QSSA solver
$\tau_{washout}$	washout time scale
υ_{il}	stoichiometric coefficient for species *i* in reaction *l*
ω_i	cloud enhancement factor (dependent on photochemical reaction)
ξ	CMAQ's generalized coordinate (positively increasing upward)
ζ	entropy density

1. INTRODUCTION

Air quality problems have serious implications to public health. In the United States, the Clean Air Act Amendments of 1990 (CAAA-90) provide a societal mandate to assess and manage air pollution levels to protect human health and the environment. National and regional policies are needed for reducing and managing the amount and type of emissions that cause acid, nutrient, and toxic pollutant deposition to ecosystems at risk and for enhancing the visual quality of the environment. Human exposure to air pollutants is one of the factors determining air pollution standards. The U.S. Environmental Protection Agency (EPA) established National Ambient Air Quality Standards (NAAQS) that require development of effective emissions control strategies for such pollutants as ozone and particulate matter (PM). Eulerian air quality models (AQMs) are used both to develop optimal emission control strategies that are both environmentally protective and cost effective and to advance atmospheric science understanding. Because of the increased reliance on AQMs for the assessment of air quality impacts on human and environmental health, rapid improvements in the parameterizations of atmospheric processes such as near surface (i.e., planetary boundary layer), clouds, radiation, aerosols, and linkage to global climate processes, as well as application to regional and local air quality forecasting are needed.

To meet both the challenges posed by the CAAA-90 and the need to address the complex relationships among pollutants, the U.S. EPA developed the Models-3 Community Multiscale Air Quality (CMAQ) system (Novak et al. [1]). Traditionally, AQMs have addressed individual pollutant issues such as urban ozone, regional acid deposition, particles, nitrogen, and toxics problems separately. However, many of the air pollutants are subjected to the same meteorology and oxidation processes. Therefore, one of the main objectives of the CMAQ system is to apply a "one atmosphere" multiscale and multi-pollutant modeling approach based mainly on the "first principles" description of the atmosphere while providing means to independently but consistently model different pollutants, if necessary, using the same tools. Needs for such third-generation comprehensive air quality modeling systems were described in Dennis et al. [2] and Byun et al. [3].

The multiscale capability is supported by the governing diffusion equation in a generalized coordinate system that handles many map projections and vertical coordinate systems, a scheme that maintains dynamic consistency with the upstream (i.e., off-line) meteorology model, a nesting

approach, and a sub-grid plume-in-grid modeling technique. The multi-pollutant (i.e., ozone, acid deposition, particulates, nitrogen loading, and toxics) capability is provided by the generalized chemistry mechanism description, general numerical solver, and comprehensive description of gaseous and aqueous chemistry and modal aerosol dynamics.

Models-3 CMAQ was released initially in 1998, and subsequently revised annually through 2003 to incorporate new developments. The science documentation of CMAQ [1] describes the state of science of the model as released in 1999 in detail. This paper provides an overview of the Models-3 CMAQ system, governing equations, and science algorithms of the system, including updates relevant to the September 2003 release of the model system. However, the main goal of the Models-3 CMAQ system is continuous development of its science through community modeling participation. Advancements in air quality model technology could not be easily shared between models due to the incompatibilities in the models. Models-3 CMAQ intends to facilitate collaborative development and linking of models for meteorology, emissions, air quality, and health effects through an open-source advanced modeling system. The modular science code structure allows continuous improvement of its science components. The system is expected to provide a common vehicle to advance environmental modeling techniques for the science and air quality management communities by incorporating developments in physical and chemical science process algorithms.

2. MODELS-3 CMAQ OVERVIEW

2.1 Evolution of the Eulerian Modeling Paradigm

To understand the community modeling paradigm the CMAQ system promotes, it is important to have some historical perspective of Eulerian air quality modeling systems. Eulerian air quality modeling started in the early 1970s from the extension of the photochemical box model and trace species dispersion model. To model urban air quality, meteorological inputs were prepared using diagnostic tools that attempt minimization of the three-dimensional divergence in the flow to avoid mass consistency problems. The chemistry mechanisms used were mostly intended to simulate daytime urban ozone evolution. Urban Airshed Model (UAM) (SAI [4]) and Caltech Air Quality Model (CIT) (McRae et al. [5]) were two early urban-scale photochemical grid models. Throughout the 1980s, several other similar modeling

systems were developed to study regional air quality issues. Examples include the Regional Oxidant Model (ROM) (Lamb [6]) for regional ozone study and the Sulfur Transport and Emissions Model (STEM/STEM II) (Carmichael et al. [7]) for regional acid deposition study. Earlier versions of the Urban-Regional Model (URM) (Kumar et al. [8]) and Comprehensive Air quality Model with Extensions (CAMx) (Environ [9]) followed the same approach although more recent versions utilized different methods for meteorological data linkage. One of the contributions of the Regional Acid Deposition Model (RADM) system (Chang et al. [10]) was that it provided a more succinct marriage between the meteorological and air quality models. RADM was a well-utilized model, if not the first public model implemented in such a way, that forced an AQM to follow the meteorological model's grid and dynamic structure. With the hydrostatic Mesoscale Model Version 4 (MM4) (Anthes and Warner [11]) linked to RADM, there was no serious dynamic representation problem because the diagnostic equation derived from the mass continuity equation was used to represent the vertical motion (omega equation). However, with the advent of the nonhydrostatic Fifth Generation Penn State University/ National Center for Atmospheric Research Mesoscale Model (MM5) system (Grell et al. [12]), a higher-level consistency between the meteorological and air quality model was needed. The SARMAP Air Quality Model (SAQM) (Chang et al. [13]) was developed from RADM because the RADM formulation diverged from the MM5 governing equation. EURAD (Hass [14]) was a re-implementation of RADM and SAQM for European applications. To maintain the dynamic and numerical consistency between meteorology and air quality, development of on-line or coupled modeling systems followed. An MM5 on-line chemistry model (Grell et al. [15]) and GATOR/MMTD (Jacobson et al. [16]) provided excellent tools for interactive science research, although the systems are not practical for retrospective application studies where meteorology is held constant and emissions scenarios change in each application.

As stated before, earlier AQMs addressed individual pollutant issues such as urban ozone, regional acid deposition, particles, nitrogen, and toxics problems separately. It has become increasingly evident that pollutant issues are inter-linked and cannot be treated in isolation. Pollutants in the atmosphere are subject to a myriad of transport processes and transformation pathways that control their composition and most of them react with hydroxyl radicals. Air pollutant concentration fields are sensitive to the type and history of the atmospheric mixtures of different chemical compounds. Thus, modeled abatement strategies of

pollutant precursors, such as volatile organic compounds (VOC) and nitrogen oxides (NO_x), to reduce ozone levels may, under a variety of conditions, cause an exacerbation of other air pollutants such as particulate matter or acidic deposition. Proper modeling of these air pollutants requires that the broad range of multiple temporal and spatial scales (hereafter, multiscale) of multiple pollutant (hereafter, multi-pollutant) interactions be considered simultaneously.

In addition to the models described above, many other modeling systems that can be applied to multiscale air quality studies exist. For example, the CHIMERE model (Vautard et al. [17]) is primarily designed to produce daily forecasts of ozone and other pollutants over Western Europe and make long-term simulations. The MATCH (Multi-scale Atmospheric Transport and Chemistry) model (Robertson et al. [18]) has been developed as a flexible transport/chemistry/deposition model for atmospheric pollutants. It is used in a range of applications from urban scale studies on ~5 km or higher horizontal resolutions to continental scale studies on acid deposition and photochemistry. Meteorological data were taken from archived output of the operational HIRLAM model [19]. THOR (Brandt et al. [20]) is an integrated weather and air pollution forecast and scenario model system for air pollution applications. It is comprised of several meteorological and air pollution models, including the Danish Eulerian model (Zlatev et al. [21]), capable of operating for applications ranging from hemispheric scale over Europe and urban background scales to urban street scale. PEGASUS (PNNL Eulerian Gas Aerosol Scalable Unified System) (Fast et al. [22]) is an atmospheric chemistry modeling system consisting of the Regional Atmospheric Modeling System (RAMS) (Pielke et al. [23]) and a chemical transport model. Many of these systems have similar capabilities but their levels of completeness vary widely.

2.2. Models-3 CMAQ Modeling System

The Models-3 CMAQ modeling system has evolved from several models operated at EPA, such as RADM, SAQM, UAM, and ROM. The system consists of three primary components (meteorology, emissions, and a chemical transport model) and several interface processors. Figure 1 illustrates the relationship between the CMAQ processors and requisite interfaces with the chemical transport model. With this structure, CMAQ retains the flexibility to substitute other emissions modeling systems and meteorological models. In the initial release, the Models-3 Emission Processing and Projection System (MEPPS) produced the emissions and MM5 provided the

meteorological fields needed for the CMAQ Chemical Transport Model (hereafter, CCTM). However, within the CMAQ paradigm, the emission processing and meteorological modeling systems can be substituted with alternative processors. The arrows in Figure 1 show the flow of data through the modeling system. A set of preprocessors provides linkage mechanisms among the meteorology, emissions, and chemistry transport modeling components. These processors include: the Emission-Chemistry Interface Processor (ECIP) that translates data from the MEPPS emission model for use in the CCTM; the Plume Dynamics Model that computes geometry of sub-grid scale Lagrangian plumes for large elevated emitters; and the Meteorology-Chemistry Interface Processor (MCIP) that translates and processes outputs from the meteorology model for the CCTM. More recently, the SMOKE (Sparse Matrix Operator Kernel Emission) modeling system (Coats and Houyoux [24], Houyoux and Vukovich [25]) has been incorporated into the Models-3 system to replace MEPPS and ECIP. Initial condition and boundary condition processors (ICON and BCON) provide concentration fields for individual chemical species at the beginning of a simulation and for the grid cells surrounding the modeling domain, respectively, and the photolysis processor (JPROC) calculates temporally varying photolysis rates.

Recognizing that many meteorological models utilize specific formulations suitable for certain scales of interest, we designed CMAQ to be capable of coupling with many different meteorological models by introducing the fully compressible governing set of equations (FCGSEs). For example, although neither MM5 nor RAMS exactly follows the FGCSEs, their predictive variables can be transformed into the required variables for the CMAQ system. The "one-atmosphere" paradigm allows building of the on-line and off-line modeling system using the same set of CMAQ codes. The CMAQ system is not a monolithic model, but rather a modeling system that allows users to build customized chemical transport models (CTMs) for air quality problems. It allows integration of multiple emission processing models, meteorological models, chemistry-transport models, and analyses of inputs and outputs. Another key design choice is the use of the Models-3 Input/Output Application Programming Interface (I/O API) (Coats et al. [26]), which is layered on top of the netCDF (Rew and Davis [27]) standard direct access data file format developed at the National Center for Atmospheric Research. The netCDF data format enables cross-platform transfer of data without conversion. The I/O API is implemented in a standard FORTRAN or C callable library for all input/output in the AQM codes. Thus I/O efficiencies can be improved

by replacing this centralized library or new data types can be handled by adding new routines to the library.

3. METEOROLOGICAL MODELING FOR AIR QUALITY

Air quality modeling should be viewed as an integral part of atmospheric modeling. It is imperative that the governing equations and computational algorithms be consistent and compatible in each component of the system for a "one-atmosphere" approach. Historically, many AQMs were designed with limited atmospheric dynamics assumptions. In the model development process, simplified sets of governing equations representing a limited range of the problem scales were adopted to enable rapid development of models. However, we believe that decisions on the dynamic assumptions and choice of coordinates should not precede the determination of the computational structure of the modeling system. To provide the scalability in describing dynamics in the CCTM, which does not solve the dynamics components by itself, a description of the atmosphere with a fully compressible governing set of equations in a generalized coordinate system is beneficial (Byun [28], [29]). By recasting input meteorological data with the variables that satisfy the governing equations in a generalized coordinate system, CMAQ can follow the dynamics and thermodynamics of the meteorological model closely, regardless of the meteorological model or its native coordinate structure.

3.1. Governing Equations for the Fully Compressible Atmosphere

In most numerical weather prediction models, temperature, pressure, and moisture variables represent the thermodynamics of the system, and their dynamic equations are often expressed in the advective form. The density is diagnosed as a byproduct of the simulation, usually through the ideal gas law. For multiscale air quality applications where strict mass conservation is required, prognostic equations for the thermodynamic variables are preferably expressed in a conservative form similar to the continuity equation. Ooyama [30] proposed the use of prognostic equations for entropy and air density in atmospheric simulations by highlighting the thermodynamic nature of pressure. Entropy is a well-defined state function of the thermodynamic variables such as pressure, temperature, and density. Therefore, entropy is a field variable that depends only on the state of the fluid. Ooyama separates dynamic and thermodynamic parameters into their primary roles. An inevitable interaction between dynamics and

thermodynamics occurs in the form of the pressure gradient force.

The foundation of the CMAQ's multiscale capability is the use of a consistent governing set of equations for meteorological and air quality modeling applications as proposed by Byun [28]. We can relate the generalized meteorological curvilinear coordinates $(\hat{x}^1, \hat{x}^2, \hat{x}^3, \hat{t})$ in a conformal map projection to the rotated earth-tangential coordinates (x, y, z, t) as:

$$\begin{cases} \hat{x}^1 = mx \\ \hat{x}^2 = my \\ \hat{x}^3 = s \\ \hat{t} = t \end{cases} \tag{1a}$$

$$\begin{cases} x = m^{-1}\hat{x}^1 \\ y = m^{-1}\hat{x}^2 \\ z = h(\hat{x}^1, \hat{x}^2, \hat{x}^3, \hat{t}) = h_{AGL}(\hat{x}^1, \hat{x}^2, \hat{x}^3, \hat{t}) + z_{sfc}(\hat{x}^1, \hat{x}^2) \\ t = \hat{t} \end{cases} \tag{1b},$$

where m is the map scale factor, s is the generalized meteorological vertical coordinate, z_{sfc} is the topographic height, h is the geometric height, and h_{AGL} represents height above the ground level (AGL). Here, s is a symbol for the generalized vertical coordinate and z is a specific Cartesian coordinate representing vertical displacement, i.e., height. In the generalized coordinate system, the contravariant and covariant wind components are represented with $\hat{v}^j$ and $\hat{v}_k$. The square root of the determinant of the metric (Jacobian) is composed of the map scale factor and vertical derivative $J_s = |\partial h / \partial s|$:

$$\sqrt{\hat{\gamma}} \equiv \left|\hat{\gamma}_{jk}\right|^{1/2} = \frac{1}{m^2}\left|\frac{\partial h}{\partial s}\right| = \frac{J_s}{m^2}, \tag{2}$$

where $\hat{\gamma}_{jk} = \frac{\partial x^i}{\partial \hat{x}^j}\frac{\partial x^i}{\partial \hat{x}^k}$. The horizontal momentum equation is given as

$$\frac{\partial \hat{\mathbf{V}}_s}{\partial t} + (\hat{\mathbf{V}}_s \bullet \nabla_s)\hat{\mathbf{V}}_s + \hat{v}^3\frac{\partial \hat{\mathbf{V}}_s}{\partial s} + f\hat{\mathbf{k}} \times \hat{\mathbf{V}}_s + m^2\left(\frac{1}{\rho}\nabla_s p + \nabla_s \Phi\right) - m^2\left(\frac{\partial s}{\partial z}\right)\left(\frac{1}{\rho}\frac{\partial p}{\partial s} + \frac{\partial \Phi}{\partial s}\right)\nabla_s h = \hat{\mathbf{F}}_s \tag{3}$$

where $\hat{\mathbf{V}}_s = \hat{v}^1\mathbf{i} + \hat{v}^2\mathbf{j}$ is the horizontal contravariant wind vector on conformal map coordinates,

$\nabla_s = \hat{\mathbf{i}}\partial/\partial\hat{x}^1\big|_s + \hat{\mathbf{j}}\partial/\partial\hat{x}^2\big|_s$, f is the Coriolis factor, $\hat{\mathbf{F}}_s$ is the horizontal forcing vector, p is atmospheric pressure, $\Phi = gh$ is the geopotential height, and ρ is air density. A prognostic equation for the vertical velocity component (w) in the Cartesian coordinates is given as:

$$\frac{\partial(\rho J_s w)}{\partial t} + m^2\nabla_s \bullet \left(\frac{\rho J_s w \mathbf{V}_z}{m}\right) + \frac{\partial(\rho J_s w \hat{v}^3)}{\partial s} + \rho J_s\left(\frac{m}{\rho}\frac{\partial p}{\partial s} + \frac{\partial \Phi}{\partial s}\right)\left(\frac{\partial s}{\partial z}\right) = \rho J_s\left(F_3 + \frac{wQ_\rho}{\rho}\right), \quad (4)$$

where $\mathbf{V}_z = \hat{\mathbf{V}}_s / m$ is the horizontal wind vector represented in the Cartesian coordinate system, and F_3 is the forcing term for the w-component. The contravariant vertical velocity component is related to the Cartesian vertical velocity with:

$$\hat{v}^3 = \frac{ds}{dt} = \frac{\partial s}{\partial t} + \mathbf{V}_z \bullet \nabla_z s + w\left(\frac{\partial s}{\partial z}\right) = \frac{\partial s}{\partial t} + \left(-\frac{1}{g}\hat{\mathbf{V}}_s \bullet \nabla_s \Phi + w\right)\left(\frac{\partial s}{\partial z}\right), \quad (5)$$

where $\nabla_z = \hat{\mathbf{i}}\partial/\partial x\big|_z + \hat{\mathbf{j}}\partial/\partial y\big|_z$.

With the ideal gas law, the thermodynamic variables (i.e., temperature, entropy, pressure gradients, and density) are diagnostically related (Byun [28]). The entropy per unit volume (entropy density), ζ is defined as

$$\zeta = \rho C_{vd} \ln\left(\frac{T}{T_{oo}}\right) - \rho R_d \ln\left(\frac{\rho}{\rho_{oo}}\right), \quad (6)$$

where T is temperature, C_{vd} is the specific heat capacity for dry air at constant volume, and R_d is the dry air gas constant. T_{oo} is temperature of the reference atmosphere at the reference pressure p_{oo} = 10^5 Pascal. The atmospheric pressure is treated as a thermodynamic variable that is fully defined by the density and entropy of the atmosphere. The conservation equations for air density, entropy density, and tracer concentrations are:

$$\frac{\partial(\rho J_s)}{\partial t} + m^2\nabla_s \bullet \left(\frac{\rho J_s \hat{\mathbf{V}}_s}{m^2}\right) + \frac{\partial(\rho J_s \hat{v}^3)}{\partial s} = J_s Q_\rho \quad (7)$$

$$\frac{\partial(\zeta J_s)}{\partial t} + m^2\nabla_s \bullet \left(\frac{\zeta J_s \hat{\mathbf{V}}_s}{m^2}\right) + \frac{\partial(\zeta J_s \hat{v}^3)}{\partial s} = J_s Q_\zeta \quad (8)$$

$$\frac{\partial(\varphi_i J_s)}{\partial t} + m^2 \nabla_s \bullet \left(\frac{\varphi_i J_s \hat{\mathbf{V}}_s}{m^2} \right) + \frac{\partial(\varphi_i J_s \hat{v}^3)}{\partial s} = J_s Q_{\varphi_i} \qquad (9)$$

where φ_i is the trace species concentration (mass per unit volume), and the Q-terms represent sources and sinks of each conservative property. Although the source term for air density (Q_ρ) should be zero in an ideal case, it is retained here to capture the possible density error originating from numerical procedures in a meteorological model. Because the error influences computations of other parameters such as vertical velocity and mass conservation, it is important to minimize propagation of the error in the system. Eqs. (3), (4), (7), (8) and (9) with additional diagnostic relations form a governing set of equations for the fully compressible atmosphere.

The generalized coordinate system allows transformations among various horizontal map projections (e.g., spherical, rectangular, Lambert, Mercator, and polar stereographic), and various vertical coordinates (e.g., pressure or geometric height). For most urban and regional applications, the equations account for the choice of horizontal map projection with the map scale factor by simply changing a few scaling parameters defining the domain boundary, map origin, and orientation. The dynamics used in meteorological models are often linked to the choice of the vertical coordinate. In those vertical coordinates that depend on atmospheric pressure, the coordinate values decrease with height. To simplify implementation of the generalized coordinate in CMAQ, without a loss of generality, we redefine the terrain-following vertical coordinate s with a positive definite coordinate $\xi = \hat{x}^3$ as

$$\hat{x}^3 = \xi = \begin{cases} s & \text{(e.g., for } \sigma_z,\ \sigma_{\bar{z}}\text{)} \\ 1 - s & \text{(e.g., for } \sigma_{\tilde{p}}, \sigma_{p_o}, \eta\text{)} \end{cases}, \qquad (10)$$

where σ_z, $\sigma_{\bar{z}}$, $\sigma_{\tilde{p}}$, σ_{p_o}, and η are vertical coordinates often used in meteorological models. Refer to Table 1 for the definitions of the vertical coordinates. This restriction simplifies physical interpretation of terms in the governing equations and eventually the computer coding of the algorithms.

3.2 Meteorological Model Data for CMAQ

Air quality models are often run many times with different scenarios to understand the effects of emissions control strategies on the pollutant concentrations using the same meteorological data. Several methods have been used to supply the meteorology data to Eulerian air quality models

such as the CCTM. Seaman [31] discusses the characteristics of three types of meteorological processors: diagnostic models, dynamic models, and dynamic models with four-dimensional data assimilation (FDDA). An off-line dynamic meteorological model with FDDA typically provides meteorological data for the CCTM. However, a successful air quality simulation requires that the key parameters in the meteorological data (e.g., wind, density, moisture variables) be consistent in the meteorological model and the CCTM. CMAQ replicates the dynamic assumptions used for the meteorological simulation on the same coordinate and grid systems. The coordinate transformation is performed implicitly using the Jacobian (calculated in a pre-processor to the CCTM) within the CCTM. CMAQ allows the use of different vertical coordinates without having to exchange science process modules describing physical parameterizations. CMAQ follows the methods in Byun [28], [29] for the implementation of the generalized coordinate.

Off-line CMAQ modeling primarily relies at present on the PSU/NCAR MM5 system (Grell et al. [12]) as the meteorological driver. MM5 is a complex, full-physics, state-of-the-science community model. MM5 Version 2 allows both the hydrostatic and non-hydrostatic dynamic options, while MM5 Version 3 supports only the non-hydrostatic dynamics. MM5 has a rich pool of physics modules for the planetary boundary layer (PBL) parameterization, radiation schemes, sub-grid cloud parameterizations, explicit moisture physics, and land-surface modeling. Different model physics parameterizations are typically applied as the horizontal grid spacing is changed. For off-line air quality modeling, one-way nesting is generally used in MM5 with multiscale FDDA, following Stauffer and Seaman [32].

CMAQ can also acquire meteorology data from the Regional Atmospheric Modeling System (RAMS) (Tremback [33], Pielke et al. [23]). RAMS is a technically comparable model to MM5, but it uses a different vertical coordinate and a different horizontal grid system, and it generates a different suite of output variables. RAMS data have been linked to CMAQ (Sugata et al. [34]) by modifying a RAMS post-processor and CMAQ's Meteorology-Chemistry Interface Processor (MCIP).

3.3 Meteorology and Chemistry Model Linkage Issues

The consistency among meteorological parameters and the way they are utilized in a CTM are crucial to the success of air quality simulations. For CMAQ, MCIP translates and processes model outputs from the meteorology model for the

CCTM. MCIP is loosely based on the meteorological pre-processor of the Regional Acid Deposition Model (RADM) (Chang et al. [10]). Because many mesoscale models were not designed specifically for air quality, and they do not output all the parameters necessary for air quality simulations, diagnostic routines are implemented in MCIP to fill the data gap. In addition, MCIP converts between coordinate systems and sets up the generalized vertical coordinate, which enables the flexibility to link various meteorological models to the CCTM. MCIP also computes dry deposition velocities for various photochemical species.

The ultimate goal within the atmospheric community is the development of a fully integrated on-line meteorological-chemical model (Seaman [35]). There have been a few successful examples of integrating meteorology and atmospheric chemistry algorithms into a single computer program (Vogel et al. [36], Grell et al. [15]), including direct coupling of MM5 with MAQSIP (Xiu et al. [37]). It should be noted, however, that there can be tradeoffs to using an on-line system. On-line systems considerably increase computer resource requirements. In addition, it is impractical to maintain the fully coupled modeling system for use in routine air-quality management activities where the same meteorology is repeatedly used for various emissions scenarios. Our goal is to develop a system that can be used for both on-line and off-line modes, depending on the objective of the air quality simulation.

The Weather Research and Forecast Model (WRF) (Dudhia et al. [38], Klemp et al. [39]),which is currently under development, is a meteorological model that is expected to have both on-line and off-line capabilities for chemistry transport modeling. WRF is being designed with the following attributes, which are desirable for a meteorological model to host both on-line and off-line capabilities:

- Achieve scaleable dynamics and thermodynamics with a fully compressible form of the governing equations and a flexible coordinate system;
- Use a unified governing set of equations that is capable of dealing with air quality as well as weather forecasting;
- Maintain cell-based mass conservation by utilizing proper state variables (i.e., density and entropy, rather than pressure and temperature) and representation of the governing equations in the conservation form rather than in the advective form; and

- Implement general physical parameterizations that are designed for a wide spectrum of spatial and temporal scales.

It is anticipated that WRF will ultimately be coupled with CMAQ in both the off-line and on-line modes as WRF development matures.

4. EMISSION MODELING

Eulerian air quality models depend on preprocessed emission data to describe primary air pollutant inputs in the atmosphere and therefore they are called emission-based models (EBMs). As such, emission modeling is one of the critical components of the Models-3 CMAQ system. Initially, CMAQ modeling depended on the Models-3 Emission Processing and Projection System (MEPPS) (Benjey and Moghari [40], Benjey et al. [41]). More recently, EPA replaced MEPPS with the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system [http://www.emc.mcnc.org/products/smoke/] for the preparation of area, mobile, point, and biogenic emission data (Houyoux and Vukovich [25]). SMOKE provides gridded, temporalized (usually hourly), and speciated emission data. SMOKE not only deals with individual chemical species but also accommodates lumped (grouped) species consistent with the gas phase chemical mechanisms contained in the CMAQ model. The matrix-based internal data structure of SMOKE allows rapid processing of emissions factor applications with simple sparse matrix multiplication operations. The general procedures for modeling of emissions with SMOKE are described briefly below, followed by more specific procedures associated with each emission source type.

4.1 Emission Inventories

Historically, many emission inventories have been compiled and used for the purpose of regulatory or scientific assessment of emissions, including spatial and temporal patterns and trends. Emission inventory data are available from various sources, often from State and local air pollution control agencies. The inventories commonly include area and point source data, and aggregated estimates of mobile source emissions. The EPA compiles the data into national annual emission inventories for the United States. The spatial extent of an inventory may vary from plant-specific emission data to data for an entire county or more.

4.2 Spatial Allocation of Emission Data

Emission inventory data must be spatially allocated to the modeling grid from the geographic units in which they are available. In the United States this usually means allocating data from the county level to the user-defined grid cells proportionately to the area of the geographic unit within each cell. It is possible to substantially improve on the representativeness of spatial assignments of emission data by overlaying spatial surrogate data appropriate to the location of different types of emission sources. For example, population census tracts and road network location data may be overlain on county and grid maps to weigh the fractions of emissions attributable to source category codes representing activities proportional to population density and motor vehicles, respectively. The surrogate data currently used include combinations of population and housing census data, political boundaries, water boundaries, road networks, ports, airports, railroads, and detailed land use data. Current spatial surrogate data are available from EPA at http://www.epa.gov/ttn/chief/emch/spatial/.

4.3 Temporal Allocation of Emission Data

Emission data that are based on annual, seasonal, weekly, or daily values may be temporally allocated to hourly data used in the air quality model. Generally, this procedure is applied to regional inventories of point- and area-source emission data. Biogenic- and mobile-source emissions may be directly modeled as hourly data for the time period of interest, rather than extracted from an inventory. Allocation of emission data from time periods greater than hourly down to hourly data is accomplished by use of source category-specific seasonal, monthly, weekly, and daily temporal allocation factors. SMOKE allocates emissions to successively more highly resolved periods (annual to seasonal, seasonal to weekly, weekly to daily). The daily values are then transformed into emission values for each hour of a typical day by using user-supplied or default temporal allocation profiles. Temporal profiles developed for the National Acid Precipitation Assessment Program (NAPAP) (Fratt et al. [42]), with some more recent supplements (Moody et al. [43]), are used currently. These are available from EPA at http://www.epa.gov/ttn/chief/emch/temporal/.

4.4 Chemical Speciation of Emission Data

Chemical transport models, including the CCTM, require that emission data be provided for either individual species or lumped species. However, an initial processing step is required because emission data are usually reported for

pollutants that are aggregates of many species, such as volatile organic compounds (VOCs) or fine particles having diameters less than 2.5 μm ($PM_{2.5}$). These aggregate pollutants must be split into their component species, or speciated. The speciation takes two forms, discrete and lumped-model. In discrete speciation, a pollutant is split into the individual chemical components. The discrete components in an emission stream are determined by a number of methods including source testing, surrogate application, and engineering knowledge of the process. In lumped speciation, individual organic species are assigned to one or more model species (groups of species) according to the chemical mechanism used. The rules for assigning the discrete compounds to the model species are mechanism dependent. Currently, three mechanisms, RADM2, CB4, and SAPRC99 (Carter [44]) are available. SMOKE also provides a speciation of fine particulate matter into elemental carbon, organic carbon, particulate sulfate, and particulate nitrate. Current information on speciation data and lumped profiles are available at http://www.epa.gov/ttn/chief/emch/ speciation/.

4.5 Modeling Emissions By Source Type

Modeling of point, area, mobile and biogenic source emissions requires use of different combinations of the spatial, temporal, and speciation capabilities of SMOKE. Point source processing addresses emissions from discrete stacks or vents. Hourly emission estimates are computed by applying factors to annual emission inventory data from source category-specific temporal allocation profiles, or by using source-specific hourly continuous emissions monitoring (CEM) data, if available. The CEM data are a subset of hourly emission data (CO, NO_x, and SO_2) derived from continuous air pollutant concentration monitors attached to components of specific facilities, usually boilers or stacks of large point sources such as electric utilities. SMOKE also calculates the plume rise and initial vertical plume spread of point source emissions to determine the vertical levels of the CCTM modeling domain into which point source emissions should be introduced.

Area emissions are attributable to diffuse sources spread over areas, such as agricultural fields, large open mining operations, forests, and/or a combination of many point sources which are too small and numerous to account for individually (such as residences). Biogenic and mobile source emissions are treated separately because they may be modeled episodically using hourly meteorological data as inputs. Biogenic emissions include natural emissions from vegetation, soils, and lightning. Gaseous biogenic emissions depend on temperature, solar radiation, and land cover type.

Hourly emission rates of isoprene, monoterpenes, other VOCs, and soil NO_x for each grid cell are estimated using the Biogenic Emission Inventory System, Version 3 (BEIS3) (Pierce et al. [45], Guenther et al. [46]). The biogenic emission processing depends on ambient temperature, total horizontal solar radiation flux, and land cover inputs. BEIS3 applies emission flux factors to specific tree genera and agricultural crop types by geographic area for biogenic emission species in accordance with the following equation:

$$E_i = \sum_j A_j E_f F_{ij}(S,T), \tag{11}$$

where E_i is the emission rate for species i (e.g., isoprene, monoterpene) in mole/sec for each cell, A_j is the area of vegetation type j (m^2) in a grid cell, E_f is the emission factor (μg per m^2 per hour), and $F_{ij}(S,T)$ is an environmental factor to account for solar radiation S and temperature T.

Emissions from roadway vehicles, aircraft, trains, shipping, and off-road mobile equipment are termed as mobile source emissions. Vehicular emissions generated from sources other than roadway activities are usually included as area sources in the emission inventories. In SMOKE, estimates of gaseous and particulate on-road mobile source emissions, and limited toxic emissions, are prepared on an hourly basis by county or specific road segments (links) for periods of interest, usually several days. Non-road mobile source gaseous, particulate, and SO_2 emissions are disaggregated from annual area source inventories as a normal part of area source emission modeling. The hourly, grid-cell specific emission factors for different vehicle and roadway classifications are modeled with the Mobile 6 model (U.S. EPA [47]).

5. GOVERNING DIFFUSION EQUATION AND STURCTURE OF CMAQ CTM

The conservation equation for the trace species, Eq. (9), describes the instantaneous state of the trace concentration in the atmosphere. Because of the stochastic nature of atmospheric motions, the equation must be averaged to form a deterministic conservation equation before it can be solved numerically.

5.1 Governing Diffusion Equation

The air density and species concentration are decomposed into mean and turbulent terms (Reynolds decomposition):

$$\rho = \overline{\rho} + \rho' \tag{12}$$

$$\varphi_i = \overline{\varphi}_i + \varphi_i' , \tag{13}$$

where $\overline{\rho}$, ρ' are the mean and turbulent components of air density and $\overline{\varphi}_i$ and φ_i' the same for concentration of trace species *i*, respectively. Some of the parameters in the conservation equations (7) through (9) are nonlinearly related to each other and, therefore, direct application of Reynolds decomposition to these parameters will introduce covariance terms that complicate the turbulence equations. Instead, we define averaged mixing ratio and its fluctuation component based on Eqs. (12) and (13):

$$\overline{q}_i \equiv \overline{\varphi}_i / \overline{\rho} \tag{14a}$$

$$q_i' \equiv \varphi_i' / \overline{\rho} \tag{14b}$$

Similarly, the average contravariant wind components and their fluctuations are defined as

$$\overline{\hat{v}^k} \equiv \overline{\rho \hat{v}^k} / \overline{\rho} \tag{15a}$$

$$(\hat{v}^k)' \equiv \hat{v}^k - \overline{\hat{v}^k} \tag{15b}$$

These definitions allow the continuity equation for the Reynolds averaged variables to keep the original conservation form as

$$\frac{\partial(\overline{\rho} J_\xi / m^2)}{\partial t} + \frac{\partial}{\partial \hat{x}^k}\left(\frac{\overline{\rho} J_\xi}{m^2}\overline{\hat{v}^k}\right) = 0 \tag{16}$$

Note that the metric tensor components (m and J_ξ) that define the coordinate transformation rules are not turbulence variables. This means that we can define the coordinates based on the Reynolds averaged quantities. The vertical grids are defined incrementally between time steps when a time-dependent vertical coordinate is used.

Decomposing velocity components using Eqs. (15 a and b), we obtain a Reynolds averaged trace species conservation equation, neglecting the molecular diffusion:

$$\underset{\text{(a)}}{\frac{\partial \varphi_i^*}{\partial t}} + \underset{\text{(b)}}{\hat{\nabla}_\xi \bullet \left[\varphi_i^* \overline{\hat{\mathbf{V}}}_\xi\right]} + \underset{\text{(c)}}{\frac{\partial(\varphi_i^* \overline{\hat{v}^3})}{\partial \hat{x}^3}}$$

$$+ \hat{\nabla}_\xi \bullet \left[\overline{\rho}\sqrt{\hat{\gamma}}\hat{\mathbf{F}}_{q_i}\right] + \frac{\partial(\overline{\rho}\sqrt{\hat{\gamma}}\hat{F}_{q_i}^3)}{\partial \hat{x}^3} \tag{17}$$

$$\underbrace{\qquad}_{(d)} \qquad \underbrace{\qquad}_{(e)}$$

$$= \sqrt{\hat{\gamma}}\, R_{\varphi_i}(\overline{\varphi}_1, ..., \overline{\varphi}_N) + \sqrt{\hat{\gamma}}\, S_{\varphi_i}$$

$$\underbrace{\qquad}_{(f)} \qquad \underbrace{\qquad}_{(g)}$$

$$+ \left.\frac{\partial(\varphi_i^*)}{\partial t}\right|_{cld} + \left.\frac{\partial(\varphi_i^*)}{\partial t}\right|_{ping} + \left.\frac{\partial(\varphi_i^*)}{\partial t}\right|_{aero}$$

$$\underbrace{\qquad}_{(h)} \qquad \underbrace{\qquad}_{(i)} \qquad \underbrace{\qquad}_{(j)}$$

where $\varphi_i^* = \sqrt{\hat{\gamma}}\,\overline{\varphi}_i = (J_\xi / m^2)\overline{\varphi}_i$, and the Reynolds turbulent flux terms are $\hat{\mathbf{F}}_{q_i} = \hat{\mathbf{i}}\hat{F}_{q_i}^1 + \hat{\mathbf{j}}\hat{F}_{q_i}^2 = \hat{\mathbf{i}}\overline{q_i'\hat{v}^{1'}} + \hat{\mathbf{j}}\overline{q_i'\hat{v}^{2'}}$, $\hat{F}_{q_i}^3 = \overline{q_i'\hat{v}^{3'}}$. They are related with the Cartesian counterpart as $\hat{F}_{q_i}^1 = mF_{q_i}^x$, $\hat{F}_{q_i}^2 = mF_{q_i}^y$, $\hat{F}_{q_i}^3 = (\frac{\partial\xi}{\partial x})F_{q_i}^x + (\frac{\partial\xi}{\partial y})F_{q_i}^y + (\frac{\partial\xi}{\partial z})F_{q_i}^z$. Eq. (17) is similar to the conservation equation in the generalized coordinates suggested by Toon et al. [48]. Terms in Eq. (17) are explicitly related to the modules corresponding to the science processes in CMAQ:

(a) time rate of change of concentration (implemented with fractional time step)
(b) horizontal advection
(c) vertical advection
(d) horizontal eddy diffusion
(e) vertical eddy diffusion
(f) production or loss from chemical reactions
(g) emissions
(h) cloud mixing and aqueous production or loss
(i) plume-in-grid process
(j) aerosol process

Note that the dry deposition process can be included in the vertical diffusion process as a flux boundary condition at the bottom of the model layer.

5.2 Modular Structure of the CMAQ CTM

CMAQ is structured to accommodate many different science process modules that provide one-atmosphere and community multiscale modeling capability. In a model with the fractional time step approach, meteorological input parameters are read in or interpolated at the so-called synchronization time step at which a set of process modules complete the concentration update before marching into the next time increment. One of the distinctive features of CMAQ as compared to other AQMs is the hierarchical functional modularity of the science processor codes. By making appropriate science modules available in the CMAQ

system, users can build a specific CTM that may include all or some of the critical science processes, such as atmospheric transport, deposition, cloud mixing, emissions, gas- and aqueous-phase chemical transformations, and aerosol dynamics and chemistry.

We define the levels of modularity in the science model based on the granularity of the modeling components. The coarsest level of modularity is the distinction between the system framework and science models. The second level is the division of science sub-models. The third level of modularity involves *classes* that include driver module, processor module, data provider module, and utility module (a collection of assisting subroutines) in the CTM. The fourth level of modularity is based on the computational functionality in a processor module, e.g., science parameterization, numerical solver, processor analysis, and input/output routines. The next meaningful modularization level is the isolation of sections of code that can benefit from machine dependent optimization. While the emissions processing and the meteorology model are modular at the second level, the CCTM achieves the fourth-level modularity by employing the operator splitting, or fractional time step, concept in the science processes. The CMAQ codes are archived based on the science process classes that are identified at the third and/or fourth modularity level. The CCTM modules are divided into classes that provide necessary functions for the module integration, science processes, and data provider modules supplying information such as photolysis rates and aerosol dry deposition velocities, and others. Refer to Table 2 for the description of the science classes and subroutines called by the CMAQ driver.

A driver module orchestrates the synchronization of numerical integration across the science processes and some input/output sequences. The driver checks data consistency for a run scenario (input files, run time, and grid/coordinate information) and calls for the initialization routines (INIT). INIT initializes pollutant initial concentrations for gaseous species in molar mixing ratio units (ppm) and aerosol species in density units ($\mu g\ m^{-3}$), the same as the output units of CMAQ. The pair of couple/decouple routines (COUPLE) are instituted to convert the concentration in density units ($\overline{\varphi}_i$) to mixing ratio for the vertical diffusion and chemistry. Because the molar mixing ratio used in chemistry is just a constant multiplication of the ratio of molecular weight of gas species to air as in $\overline{m}_i = \overline{q}_i (M_{air} / M_i)$, no further conversion is necessary. These unit conversion calls limit the interchange between groups of process modules using different concentration units. The driver structure of the current CCTM is given in Figure 2. A synchronization time

step (Δt_{sync}) is used to ensure the global stability of the numerical integration at the advection time step, which is based on a Courant number limit. Nesting requires finer synchronization time steps for the fine grid domain. In the CCTM, the process synchronization time steps are represented as integer seconds due to a limitation of Models-3 I/O API that can only handle integer seconds for I/O data. All the needed data are appropriately interpolated based on the synchronization time step. For maintaining numerical stability and for other reasons, an individual process module may have its own internal time steps. In general, each science process module uses the synchronization time step as the input time step of required environmental data. The global output time steps can be set differently from the synchronization time step. Usually, the output time step is set to be one hour, but sub-hourly output down to the synchronization time step is possible.

Science process classes in the CCTM are identified in Eq. (17). The only data dependencies among the CCTM science modules are the trace species concentration fields (i.e., concentrations are the objects of science process operations) and the model integration time step. To facilitate modularity and to minimize data dependency in the CCTM, we store concentrations in global memory while the environmental input data are obtained from random-access files and interpolated to the appropriate computational (synchronization) time step. This realizes the characteristic "thin-interface" structure of the CCTM:

- Concentration and timing data are the only arguments of module driver routines.
- Environmental data are provided with a standard I/O interface.
- Model structure (i.e., grid and chemical mechanism) data are passed through the shared include files.
- Standard physical constants are given in fixed shared include files.

5.3 Grid and Nesting

CMAQ uses a regular (structured and uniform) horizontal grid system for simplicity. To ensure accuracy of simulations for desired areas at fine scales, a static grid nesting technique is used. It involves the sequential placement of multiple finer-scale meshes in desired regions of the domain so as to provide increased spatial resolution locally. The spatial resolution of the coarse grid is usually an integer multiple of that of the fine grid. First, the coarse grid solution is marched forward one time step. This solution provides boundary and initial conditions (both

concentration and flux) to the fine grid solution that is advanced at a smaller time step (usually an integer fraction of the time step of coarse grid). It is customary to set the time step ratio equal to the grid size ratio to retain the numerical accuracy at the same order of approximation. When the fine grid solution catches up with the coarse grid solution, the former may or may not be used to update the latter (i.e., two-way vs. one-way nesting). There are a few shortcomings of using grid nesting. One is the tendency for propagating dispersive waves to discontinuously change their speeds upon passing from a coarse mesh to the next finer mesh and to reflect off the boundaries of each nest due to the mismatch across the mesh boundaries. CMAQ includes a Lagrangian plume-in-grid (PinG) method to resolve the spatial scale of plumes emanating from major elevated point source emitters.

5.4 Process Analysis

The CMAQ CTM predicts the spatial and temporal distributions of ambient air pollutant concentrations affected by important physical and chemical processes as described above. Because CMAQ uses the fractional time step method, the species concentrations are determined from successive changes in concentrations due to the atmospheric processes. In turn, changes in concentrations by each process can be accumulated to account for the relative contribution of the atmospheric process determining the pollutant concentrations. This is the so-called process analysis output. This type of information has been used to assess how a model generates its predictions and to identify potential sources of error in the model formulation (Pleim [49], Jeffries and Tonnesen [50], Jang et al. [51], [52]). Process analysis can be described in two parts: integrated process rate (IPR) analysis for all the atmospheric processes in the model and integrated reaction rate (IRR) analysis, which provides detailed information specific to the chemical reactions. IRR analysis deals with the details of the chemical transformations that are described in the model's chemical mechanism. This is particularly useful for investigating mechanistic differences under different chemical regimes (e.g., VOC- versus NO_x-limited conditions). Either analysis method can be applied independently of the other. The current CMAQ's implementation of IRR analysis addresses only gas-phase reactions. Nevertheless, the concepts should be adaptable to the modules simulating aerosol formation and aqueous chemistry as well. This is an area for future enhancement in the CCTM.

6. ATMOSPHERIC TRANSPORT PROCESSES

6.1 Advection Processes: HADV, VADV and ADJCON

The advection process relies on the mass conservation characteristics of the continuity equation,

$$\frac{\partial \varphi_i^*}{\partial t} = -\nabla_\xi \bullet \left(\varphi_i^* \overline{\hat{\mathbf{V}}}_\xi \right) - \frac{\partial (\varphi_i^* \overline{\hat{v}^3})}{\partial \xi} . \qquad (18)$$

Given dynamically and thermodynamically consistent meteorological data, we can maintain conservation of both the trace species mixing ratio and mass at the model synchronization time step. When a CTM is coupled with a meteorology model that does not satisfy the necessary characteristics for mass conservation (described earlier) and the meteorological data and the numerical advection algorithms are not exactly mass consistent, only conservation of mixing ratio, not the tracer mass, should be expected. In such a case, we need to solve a modified advection equation in CTMs (Byun [29]):

$$\frac{\partial \varphi_i^*}{\partial t} = -\nabla_\xi \bullet \left(\varphi_i^* \overline{\hat{\mathbf{V}}}_\xi \right) - \frac{\partial (\varphi_i^* \overline{\hat{v}^3})}{\partial \xi} + \varphi_i^* \frac{Q_\rho}{\rho} , \qquad (19)$$

where Q_ρ is the mass consistency error term. Eq. (19) shows that the correction term has the same form as a first-order chemical reaction whose reaction rate is determined by the mass consistency error (normalized with air density) in the meteorology data. In CMAQ the advection process is divided into horizontal (HADV) and vertical (VADV) components. This distinction is possible because the mean atmospheric motion is mostly in horizontal planes while the vertical motion is related with the interaction of dynamics and thermodynamics. Horizontal advection is further decomposed into x- and y-direction processes, where the solution alternates between (x, then y) and (y, then x) directions to maintain a symmetric form of the advection process calls.

The current CCTM codes include two flux-based advection algorithms, the piecewise parabolic method (PPM) (Colella and Woodward [53]) and the Bott scheme (Bott [54]). A detailed test result of these schemes in comparison with other advection algorithms is reported in Odman [55]. In the PPM, the concentration distribution is assumed to be parabolic in any given grid cell. The PPM is absolutely positive definite and monotonic, while the Bott scheme is non-monotonic and could generate unwanted local extreme values for trace species, such as aerosol number density, that have large concentration gradients. PPM is somewhat more

diffusive than the Bott scheme, but can handle transport of species with large spatial inhomogeneity and dynamic range better because of the monotonic character of the algorithm.

The mass conservation characteristic of Eq. (19) is heavily dependent on the quality of the wind data provided. Dynamic assumptions of the meteorological model and the application method of its numerical advection algorithms (e.g., time splitting of 3-D advection into a sequence of 1-D operators) influence the mass conservation characteristics of the model. They can be a source of error in the species continuity equation even in the absence of emissions. The modules in the ADJCON class fix this operational error by solving the following equation:

$$\frac{\partial \varphi_i^*}{\partial t} = \varphi_i^* \frac{Q_\rho}{\rho} . \tag{20}$$

Here we want to emphasize that the artificial distinction of advection modules between horizontal and vertical processes is not adequate and that all three modules (HADV, VADV, and ADJCON) should be considered as an integral unit for solving the physical advection process of trace species.

6.2 Diffusion Processes

Turbulent diffusion only is modeled in the CCTM. The concentration units used for horizontal and vertical turbulent diffusion processes (i.e., HDIFF and VDIFF) are density (φ_i) and molar mixing ratio (m_i), respectively. We have chosen m_i as the generic concentration unit for the vertical diffusion to coordinate with the emissions units in the data. Because the ratios of molecular weights relative to air are constant, equations for the vertical diffusion in molar mixing ratio are mathematically equivalent to those in the mass mixing ratio, q_i.

Horizontal diffusion

When the turbulent flux terms are parameterized via the gradient transport theory (K-theory), the horizontal diffusion equation is given as (Byun et al. [56]):

$$\left.\frac{\partial \varphi_i^*}{\partial t}\right|_{hdiff} = \frac{\partial}{\partial \bar{x}^1}\left[\hat{K}_*^{11} \frac{\partial \bar{q}_i}{\partial \bar{x}^1}\right] + \frac{\partial}{\partial \bar{x}^2}\left[\hat{K}_*^{22} \frac{\partial \bar{q}_i}{\partial \bar{x}^2}\right], \tag{21}$$

where $\hat{K}_*^{11}$ and $\hat{K}_*^{22}$ are the horizontal components of the contravariant eddy diffusivity multiplied with factor $\sqrt{\hat{\gamma}}\bar{\rho}$.

The off-diagonal terms are neglected as a practical simplification. Because the meteorological input parameters are read in or interpolated at each Δt_{sync}, we approximate $\sqrt{\hat{\gamma}\bar{\rho}}$ to be constant for the duration of a synchronization time-step for integrating the diffusion process with the fractional time-step method. The horizontal diffusion equation is solved with an explicit finite difference method with a sufficiently small time step to ensure positive solutions. The reason an implicit scheme is not used here is to minimize the computer memory requirement for handling large horizontal grids in the subroutine. For the explicit scheme, the time-step should be chosen such that numerical stability and positivity of the solutions are maintained. With an appropriate Courant number for the horizontal diffusion, β_{hdiff}, the time-step can be determined with:

$$\Delta t_{hdiff} = \beta_{hdiff} \frac{(\Delta x)^2}{\hat{K}_H^{\max}} . \qquad (22)$$

where $\hat{K}_H^{\max}$ is the domain maximum contravariant horizontal diffusivity and β_{hdiff} = 0.75 in the CCTM. The contravariant eddy diffusivity is related to the Cartesian counterpart as $\hat{K}^{11} = \hat{K}^{22} = mK_H$. The K_H formulation used in CMAQ is described below.

Horizontal eddy diffusivity

The diffusion process must represent the effects of physical diffusion on pollutant dispersion. Although our understanding of horizontal turbulence is rather limited, appropriate accounting of physically-based horizontal diffusion is necessary. We can identify certain types of nonphysical horizontal diffusion, such as the numerical diffusion resulting from the inconsistency (i.e., errors in higher-order expansion terms) in the advection scheme and artificial diffusion resulting from the instantaneous dilution of emissions and concentrations by the finite volume of the Eulerian grid cells. Ideally, the driver meteorology model should provide the eddy diffusivity for use in the CCTM, and therefore, whenever appropriate, a pass-through approach is preferred. However, the horizontal diffusion process is sometimes omitted in meteorology models because the numerical diffusion associated with the advection was already large. The eddy diffusivity must be estimated in the CCTM when it is not passed through from the meteorology model.

The specified horizontal diffusion term in Eulerian dispersion models, when combined with the effects of the input wind fields, the numerical diffusion of the advection

scheme and the instantaneous dilution in the grid cell should accurately simulate the diffusion that is observed in the atmosphere. The eddy diffusivity must take into account the genuinely advective characteristics of wind flows. For example, Smagorinsky's [57] horizontal diffusivity algorithm accounts for the transport (stretching and shearing deformation) characteristics of wind flows:

$$K_{HT} = 2\alpha_o^2 (S_\Gamma^2 + S_\Lambda^2)^{1/2} (\Delta x)^2, \tag{23}$$

where $\alpha_o \cong 0.28$ and stretching strength (S_Γ) and shearing (S_Λ) strength are defined by

$$S_\Gamma = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) \tag{24a}$$

$$S_\Lambda = \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) \tag{24b}$$

Because Eq. (23) relies on the grid-scale wind components, it is not suitable for estimating the sub-grid scale diffusion not resolved by the modeled wind fields. Furthermore, for a coarse resolution where numerical diffusion is already large, use of this formula seems inadequate. For simulations with a larger grid size, eddy diffusivity may be parameterized to counteract the numerical diffusion (e.g., Byun et al. [56]):

$$K_{HN}(\Delta x) = K_{Hf}(\Delta x_f)\left(\frac{\Delta x_f}{\Delta x}\right)^2, \tag{25}$$

where $K_{Hf}(\Delta x_f)$ stands for a uniform eddy diffusivity at a fixed resolution Δx_f. In CMAQ $K_{Hf}\big|_{\Delta x_f = 4km} = 2000$ ($m^2 s^{-1}$) is used. The formula, however, is inadequate for a very fine grid size where the physical dispersion dominates over the numerical diffusion.

The difference in the grid size dependency between Eq. (23) and Eq. (25) is striking. A heuristic method combining the two formulae is suggested here as an analogy to the resistance law concept used for the estimation of deposition velocity:

$$\frac{1}{K_H} = \frac{1}{K_{HT}} + \frac{1}{K_{HN}}. \tag{26}$$

This formula attempts to resolve the dichotomy existing between the contrasting dependency on grid resolution in the components of horizontal diffusivity. Figure 3 evaluates Eq. (26) for various grid sizes and magnitudes of deformation. For a large grid size, the effect of the transportive dispersion

is minimized while for a small grid size the impact of the numerical diffusion term is reduced.

Vertical Diffusion

Sub-grid scale vertical diffusion of trace pollutants in the atmospheric boundary layer is another one of the important physical processes. Emissions can be included either in the vertical diffusion or gas-phase chemistry module. The vertical diffusion equation is written in terms of the mixing ratio as follows:

$$\left.\frac{\partial \bar{q}_i}{\partial t}\right|_{vdiff} = -\frac{\partial(\hat{F}^3_{q_i})}{\partial \hat{x}^3} + \left(\frac{Q_{\varphi_i}}{\bar{\rho}}\right) - \hat{F}^3_{q_i}\frac{\partial\left[\ln(\sqrt{\hat{\gamma}}\bar{\rho})\right]}{\partial \hat{x}^3}. \qquad (27)$$

The last term in Eq. (27) represents the coordinate divergence term. The importance of the vertical gradient of $\sqrt{\hat{\gamma}}\bar{\rho}$ on the turbulence flux depends on the type of vertical coordinates: the term vanishes for a mass-conserving hydrostatic pressure coordinate, while it does not for a fixed height coordinate. In any case, this effect is neglected in the CMAQ. This formulation in the Reynolds flux term can be represented with either local- or nonlocal-mixing parameterization schemes. Local closure assumes that turbulence is analogous to molecular diffusion, i.e., the flux at any point in space is parameterized by known mean values at the same point (Stull [58]).

The current CMAQ system includes vertical diffusion modules with the *K*-theory and a simple non-local closure scheme. Because of its simplicity, the *K*-theory is widely used in both meteorological and air quality models. The resulting diffusion equation is discretized to form a tridiagonal system, which is solved with a semi-implicit method based on a Thomas algorithm [59](Gaussian elimination without pivoting, followed by back substitution). In atmospheric models, a nonlocal scheme is suggested to be used in the presence of convective conditions where eddies are larger than the grid size and the *K*-theory fails to represent vertical mixing adequately. Therefore, the Asymmetric Convective Model (ACM) (Pleim and Chang [60]), a non-local closure model for the convective boundary layer, is also available. The ACM was originally developed for use in RADM and has also been applied to SAQM (Chang et al. [61]) and MM5.

Vertical eddy diffusivity

Again, the driver meteorology model ideally should provide the vertical eddy diffusivity for use in a CTM and therefore, whenever appropriate, a pass-through approach is preferred.

To provide a default ability to link to generic meteorological data, an eddy diffusivity parameterization is provided with the CCTM. The contravariant vertical eddy diffusivity is related to the components of the diffusivity tensor in the Cartesian coordinates as:

$$\hat{K}^{33} = J_{\xi}^{-2}\left[\left((m\frac{\partial h}{\partial x^1})^2 + (m\frac{\partial h}{\partial x^2})^2\right)K_H + K_{zz}\right] \qquad (28)$$

where K_{zz} is the eddy diffusivity in the geometric height coordinate. One of the problems with first-order closure is finding a rational basis for the eddy diffusivity parameterization. Only routinely measured or model-resolvable meteorological variables are used to explicitly specify a *K*-profile. Except for the cases with complex topography the effects of the horizontal gradient terms in Eq. (28) are small, and therefore, neglected. The set of eddy diffusivity formulations used in CMAQ is an extension of those used in RADM, which is a vertically integrated version of Brost and Wyngaard [62]. Refer to Chang et al. [10], Hass et al. [63], and Byun and Dennis [64] for the details.

Dry deposition

In Eulerian AQMs, the deposition process affects the concentrations in the lowest layer as a boundary flux condition. Considering the deposition process as the diffusion flux at the bottom of the model, we can relate the boundary condition in the generalized coordinate system to that of the Cartesian coordinate system as

$$\begin{aligned}\hat{F}^3_{q_i}\Big|_{dep} &= (\frac{\partial \xi}{\partial x})F^x_{q_i}\Big|_{dep} + (\frac{\partial \xi}{\partial y})F^y_{q_i}\Big|_{dep} + (\frac{\partial \xi}{\partial z})F^z_{q_i}\Big|_{dep} \\ &= (\frac{\partial \xi}{\partial z})F^z_{q_i}\Big|_{dep}\end{aligned} \qquad (29)$$

where side-wall deposition flux terms $F^x_{q_i}\Big|_{dep}$ and $F^y_{q_i}\Big|_{dep}$ are negligible. Then, the effects of dry deposition on the species concentration is accounted for by the following relationship:

$$\left.\frac{\partial \varphi_i^*}{\partial t}\right|_{dep} \approx -\frac{v_d}{h_{dep}}\varphi_i^*\Big|_{layer\,1} \qquad (30)$$

where h_{dep} is the thickness of the lowest model layer in the geometric height coordinate. In the derivation of Eq. (30), we assume that the deposition flux is constant in the lower part of the surface layer (i.e., a constant flux layer). In CMAQ, a dry deposition estimation method from Wesely [65] is available as a default option. Also a different

Models-3/CMAQ approach by Pleim et al. [66] is available for the special case when the aerodynamic resistance and the canopy or bulk stomatal resistance are obtained from MM5 with a Pleim/Xiu PBL/land-surface option. Details of both methods are described in Byun et al. [67].

7. ATMOSPHERIC CHEMISTRY PROCESSES AND SOLVERS

The representation of chemical interactions among atmospheric constituents is an essential element of the AQMs. Although many atmospheric pollutants and their precursors are emitted as gases and react primarily in the gaseous phase, environmental problems such as acid deposition and the formation of aerosols also involve pollutant interactions among the gas, liquid and solid phases. Thus, a multi-pollutant, multi-phase, one-atmosphere modeling approach is often needed for a comprehensive representation of chemical processes taking place in the atmosphere.

The CMAQ modeling system was designed to facilitate such an approach. For purposes of computational efficiency, however, the treatments of gas-phase chemistry, aqueous chemistry, and heterogeneous processes involving the formation of aerosols are divided into three separate modules. This section describes the treatment of gas-phase chemistry and the linkages added to simulate the interactions of gas-phase constituents with those in other phases. The methods used to simulate aqueous chemistry and aerosol formation are described in subsequent sections.

7.1 Photolysis Rates

Simulation accuracy of the CMAQ chemical system is highly dependent upon the accuracy of photolysis rates, which influence the primary sources of radicals in the troposphere. The photolysis rate (*J*-value) for a photodissociation reaction (*i*) is computed by

$$J_i = \int_{\lambda_1}^{\lambda_2} F(\lambda)\sigma_i(\lambda)\phi_i(\lambda)d\lambda \qquad (31)$$

where J_i (min^{-1}) is the photolysis rate, $F(\lambda)$ the actinic flux (photons cm^{-2} min^{-1} nm^{-1}), $\sigma_i(\lambda)$ the absorption cross section for the molecule undergoing photodissociation (cm^2 $molecule^{-1}$), $\phi_i(\lambda)$ the quantum yield of the photolysis reaction (molecules $photon^{-1}$), and λ the wavelength (nm). Absorption cross sections and quantum yields are functions of wavelength, and may also be functions of temperature

and pressure. They are usually determined by laboratory experiments for specific photolytic species. Actinic flux measures the spectral radiance integrated over all solid angles per unit area on the spherical receiving surface (compare this with irradiance, which is the radiance falling on a horizontal surface). Thus, the actinic flux can be called spherical spectral irradiance. The actinic flux changes with time of day, longitude, latitude, altitude, and season, and is governed by the astronomical and geometric relationships between the sun and the earth. It is affected by the earth's surface albedo as well as by atmospheric scattering and absorption.

The current approach for setting photolysis rates in CMAQ follows the approach used in RADM (Chang et al. [10]). The photolysis rates are estimated in two processing stages: first, a table of clear-sky photolysis rates is calculated for specified heights, latitudes, and hours from local noon; and then photolysis rates are interpolated from the table within the CCTM based on grid cell location and the model time, and are corrected for cloud cover. This approach is computationally efficient and has been shown by Madronich [68] to give clear-sky photolysis rates within the uncertainty of the surface-based measurements.

Calculation of clear-sky photolysis rate table

A preprocessor (JPROC) calculates clear-sky photolysis rates for 6 latitudinal bands (at every 10 degrees for 10°- 60° N), 7 altitudes (0 km, 1 km, 2 km, 3 km, 4 km, 5 km, and 10 km), and ±9 hours from local noon (0–8 h). The delta-Eddington two-stream radiative transfer model (Joseph et al. [69], Toon et al. [48]) is used for computing the actinic flux. The two-stream approximations are limited in application to cases where the scatter is not highly anisotropic. In computing the actinic flux, a description of the extraterrestrial radiation, aerosol, ozone absorption, oxygen absorption in the Schumann-Runge Bands, Rayleigh scattering (WMO [70]) and surface albedo are provided to the radiation model. Users can specify the extraterrestrial radiation; however, these data should be resolved at wavelengths that capture the features important to the photolysis reactions of interest. A modified WMO extraterrestrial radiation data distribution (Chang et al. [71]) is used as input to JPROC, which has a variable wavelength resolution ranging from 1 to 10 nm.

Users can specify the absorption cross section and quantum yield data. As default, CMAQ contains standard sets of cross section/quantum yield data for the CB4, the RADM2, and the SAPRC mechanisms in JPROC. The O_2 and O_3 absorption cross section data are from NASA data (DeMore

et al. [72]) and seasonal vertical ozone profiles are used. If total ozone column data are available (such as data measured by the Total Ozone Mapping Spectrometer (TOMS) instrument aboard the sun-synchronous, polar-orbiting Nimbus satellite), then the interpolated profiles are uniformly re-scaled to match the measured total ozone column data. TOMS data are archived and available at the National Satellite Service Data Center (NSSDC) in the form of digital daily maps with a resolution of 1 degree latitude by 1.25 degrees longitude. The TOMS data are averaged over each latitudinal band in JPROC. To account for the effects of the temperature and pressure on the absorption cross sections and quantum yields for each photolysis reaction, JPROC utilizes seasonal vertical profiles of temperature and pressure as in RADM.

The albedo data by Demerjian et al. [73], given as a function of wavelength, are used in the current version of JPROC. A single vertical profile of aerosol attenuation coefficients (Elterman [74]) is used in JPROC. A future CMAQ version with two-way interactions with a meteorological modeling system will allow calculations of radiative fluxes and photolysis rates with predicted aerosol parameters from CMAQ.

Interpolation and cloud attenuation

The reaction-dependent photolysis rates for each cell are estimated by the interpolation of the clear-sky values from the table based on the latitude, height, and time from local noon. Effects of clouds on the photolysis rates are estimated following RADM. They depend on the relative location (below, above, or within) of the cell to the cloud. The below cloud photolysis rate (J_{below}) is calculated as:

$$J_{below} = J_{clear}\left[1 + f_c(1.6t_r \cos\theta - 1)\right], \tag{32}$$

where f_c is the cloud coverage fraction (cloud fraction is interpolated from hourly data for each grid cell), θ is the zenith angle, and t_r is the cloud transmissivity. Below cloud photolysis rates will be lower than the clear-sky values due to the reduced transmission of radiation through the cloud. The cloud transmissivity is calculated by:

$$t_r = \frac{5 - e^{-\delta_{cld}}}{4 + 3\delta_{cld}(1 - f_{sc})} \tag{33}$$

where f_{sc} is the scattering phase function asymmetry factor (assumed to be 0.86) and δ_{cld} is the cloud optical depth. We have replaced the cloud optical depth equation in RADM with one taken from Stephens [75]. The original

RADM required an estimate of the cloud droplet radius, which was assumed to be 10 μm. The empirical formula for δ_{cld} from Stephens [75]:

$$\log(\delta_{cld}) = 0.2633 + 1.7905 \ln[\log(W)] \quad (34)$$

is only a function of liquid water path (W), where $W = L\Delta z_{cld}$ (g m^{-2}), L is the liquid water content (g m^{-3}), and Δz_{cld} is the cloud thickness. The above cloud top photolysis rates are calculated as:

$$J_{above} = J_{clear}\left[1 + f_c\omega_i(1 - t_r)\cos\theta\right] \quad (35)$$

This equation allows enhancement of photolysis rates above the cloud due to the reflected radiation from the cloud. It also includes a reaction-dependent coefficient ω_i, which allows further enhancements above the cloud top (Chang et al. [10]). Within the cloud, the cloud correction factor is a simple linear interpolation of the below cloud factor at cloud base to the above cloud factor at cloud top. Once computed, the below, above, and within cloud factors are used to scale the clear sky photolysis rates to account for the presence of clouds. In the current implementation, all cloud types (including clouds composed of ice crystals) are treated the same way as described above.

7.2 Gas-phase Chemistry Mechanisms

Chemical reactions taking place in the gas-phase are represented in AQMs by means of chemical kinetic mechanisms. At present, the CMAQ system includes three basic gas-phase chemical mechanisms that were originally developed to address issues associated with urban and regional scale ozone formation and acidic deposition—the CB4 [Gery et al. [76], RADM2 [Stockwell et al. [77] and SAPRC99 [44] mechanisms. Two variants of the RADM2 mechanism that have a more up-to-date and comprehensive treatment of isoprene chemistry are also included.

To facilitate using multiple chemical mechanisms, the CMAQ system has been equipped with a generalized chemical mechanism processor and two generalized gas-phase chemistry solvers--the Sparse Matrix Vectorized Gear (SMVGEAR) and a variant of the Quasi Steady-State Approximation (QSSA) approach. The generalized mechanism processor and solvers allow existing mechanisms to be modified (or a completely new mechanism to be added) without having to make code changes to the CMAQ CTM. In addition to the generalized solvers, the CMAQ system includes two mechanism-specific solvers - the Euler Backward Iterative (EBI) and a Modified

Euler Backward Iterative (MEBI) - that are faster than either of the generalized solvers and more accurate than the QSSA solver. Thus far, the fastest solver, EBI, has been developed only for the CB4 mechanism, but MEBI solvers are available for other mechanisms.

Brief descriptions of the basic gas-phase mechanisms currently in the CMAQ system are included below. It should be noted that, methane is assumed to have a constant mixing ratio of 1.85 ppm in CMAQ. Hence, the second-order reaction of the hydroxyl radical with methane in each mechanism is replaced by a pseudo first-order reaction using the assumed methane abundance. Other changes made to the mechanisms include extensions to isoprene chemistry and extensions to provide linkages to aqueous chemistry and aerosols. With respect to linkages to heterogeneous processes, four different versions of each mechanism have been created – one with no linkages (for simulating gas-phase chemistry only), one with linkages for modeling gas-phase and aqueous chemistry only, one with linkages for modeling gas-phase chemistry and aerosols only, and one with linkages for modeling all three processes collectively. While it is anticipated that the fully coupled mechanisms would be used most often, the other less detailed treatments are included as options to provide savings in computer resources if the user desires to omit some of the heterogeneous processes.

CB4 mechanism

The CB4 mechanism is a lumped structure type that is the fourth in a series of carbon-bond mechanisms, and differs from its predecessors notably in the detail of the organic compound representation. The version of the core CB4 mechanism used in the CMAQ system includes 36 species and 93 reactions of which 12 are photolytic. It is based on Gery et al. [76], but also includes several changes that have been made to the CB4 since the original publication. In 1991, the rate constants for the formation and decomposition of PAN were changed, and a termination reaction between the XO2 operator and the HO2 radical was added. Subsequently, termination reactions for the XO2N operator were also added. An updated isoprene chemistry mechanism based on the work of Carter and Atkinson [80, 81] was developed and incorporated in CB4 in 1996. The conversion of gaseous sulfur dioxide to gaseous sulfuric acid has also been added to the original mechanism. Finally, the original CB4 mechanism used simple Arrhenius Law rate constant forms that were derived from more complex temperature- and pressure-dependent rate constant expressions for some rate constants. Because the top of the CMAQ domain usually extends to a height where pressure

dependency becomes important, the CMAQ version of CB4 uses pressure-dependent forms.

RADM2 mechanism

RADM2 is a lumped species mechanism that uses a reactivity-based weighting scheme to account for lumping chemical compounds into surrogate species [77]. The base mechanism implemented in the CMAQ system contains 57 model species and 158 reactions, 21 of which are photolytic. Other than the methane reaction change, only one other change was made to the mechanism as described in the original publication. The reaction of the hydroxyl radical with CSL (cresol and other hydroxyl substituted aromatics) was restructured to eliminate the negative product coefficient. This change does not alter the gas-phase chemistry predictions.

SAPRC99 mechanism

The SAPRC99 chemical mechanism included in CMAQ is a condensed, fixed parameter version of a detailed variable lumped parameter mechanism [44]. The base mechanism contains 72 model species and 214 reactions, of which 30 are photolytic. The parameters for the lumped species were derived from the ambient mixture data used in the reactivity simulations of Carter [78,79]. The only non-methane related change made to the condensed mechanism described in reference [44] is the addition of the reactions of formic, acetic and higher organic acids with the hydroxyl radical. These reactions were taken from the detailed SAPRC99 mechanism that forms the basis for the condensed CMAQ version.

Isoprene extensions

Over the past few years, the importance of isoprene in contributing to the ozone formation in the troposphere has become more fully understood. Its representation in earlier gas-phase mechanisms was substantially condensed, partially because of computational resource considerations and partially due to uncertainties about its degradation products and their subsequent reaction pathways. Results of recent mechanistic and environmental chamber studies have led to a greater understanding of isoprene chemistry, and thus to improved mechanism representations [80]. In the CMAQ system, two variants of the RADM2 mechanism have been created that have the original RADM2 isoprene chemistry replaced with different levels of more detailed isoprene chemistry. The two levels of detail are referred to as the one-product and the four-product Carter isoprene mechanisms [81]. Both are condensed from a more detailed

isoprene mechanism. The one-product form lumps the major products of isoprene degradation into a single species, whereas the four-product mechanism represents major isoprene degradation products by means of four products – methacrolein, methyl vinyl ketone, methacrolein's PAN analogue, and other lumped products. The CB4 isoprene chemistry in CMAQ is an analogue of Carter's one-product form, whereas the SAPRC99 mechanism in CMAQ incorporates the four-product version.

Aerosol extensions

Secondary atmospheric aerosols can be formed from both inorganic and organic gaseous precursors. The major inorganic precursors are H_2SO_4 vapor formed from SO_2 and HNO_3 formed primarily from reactions involving NO_2 and N_2O_5. All of the mechanisms in CMAQ include gas-phase chemical pathways for forming these inorganic aerosol precursors. As discussed in section 8, CMAQ also includes a heterogeneous pathway for converting N_2O_5 to HNO_3. Since the N_2O_5 heterogeneous pathway is expected to predominate over the homogeneous gas-phase hydrolysis reaction [82], the gas-phase hydrolysis reaction is omitted whenever CMAQ is configured to include aerosol formation. When aerosol formation is not simulated, the gas-phase hydrolysis reaction is included. The formation of secondary organic aerosols (SOA) in the atmosphere occurs via the oxidation of organic compounds to form semi-volatile gases that can subsequently condense on particles. In CMAQ, the semi-volatiles can be formed via the oxidation of five primary organic groups (alkanes, cresol, high-yield aromatics, low-yield aromatics, and terpenes). At present, production of SOA precursors from alkanes in the CB4 mechanism is omitted since the CB4 mechanism species representing alkanes includes fragments of many other classes of compounds. Methods for including SOA production from alkanes in CB4 will be investigated for future versions of CMAQ. Tests conducted with the other CMAQ mechanisms indicate that the alkane pathway is relatively small under most conditions. Finally, the oxidation of terpenes is an important pathway for the production of SOA. The SAPRC99 mechanism includes an explicit terpene species, but in RADM2 and CB4, terpenes are lumped with other organic compounds. Hence, a special set of reactions was added to these two mechanisms to track the loss of terpenes due to reactions with O_3, OH, and NO_3. In these reactions, the oxidants are regenerated so that that there is no impact on the overall chemistry.

Aqueous chemistry extensions

As described in section 8.2, the aqueous chemistry module incorporated in CMAQ is derived from the diagnostic cloud model used in RADM version 2.6. No modifications were required to link either the RADM2 or the SAPRC99 gas-phase mechanisms to the aqueous chemistry model. To improve the CB4 gas-phase linkage to aqueous chemistry, four new species were added. First, three organic acid products that had been omitted in the original mechanism were added back in - formic acid, acetic acid, and peroxyacetic acid. Second, a species named MHP was added to represent methylhydroperoxide. In the CMAQ version of CB4, MHP is formed via the reaction of an XO2 operator that does not strictly represent the methylperoxy radical, and no decomposition reactions for MHP are included. Hence, MHP represents a potential upper limit to the concentration of methylhydroperoxide. Since all four of the new species are products and they do not react with any other mechanism species, their inclusion does not affect gas-phase chemistry.

7.3 Gas-Phase Chemistry Solver

The mass conservation equation representing gas-phase reactions is given as

$$\left.\frac{\partial \sqrt{\hat{\gamma}}\overline{\varphi}_i}{\partial t}\right|_{chem} = \sqrt{\hat{\gamma}} R_{\varphi_i}(\overline{\varphi}_1, \overline{\varphi}_2, ..., \overline{\varphi}_N) + \sqrt{\hat{\gamma}} S_{\varphi_i}, \quad (36)$$

where R_{φ_i} and Q_{φ_i} represent chemistry reactions and source terms, respectively, and N is the total number of species in the chemical mechanism. Although various units could be used in the solution of gas-phase chemistry problems, the use of parts per million by volume is convenient for numerical reasons since the magnitude of the values for trace gases is small. Because the computational grid is constant for the duration of a synchronization time step, the Jacobian and species density in $\sqrt{\hat{\gamma}}\overline{\varphi}_i$ of Eq. (36) can be decoupled, and the density of a trace gas species converted to volumetric (or molar) mixing ratio units. This leads to the following conservation equation for gas-phase chemistry in terms of the time-rate of change of the volumetric mixing ratio for each species:

$$\left.\frac{\partial \overline{m}_i}{\partial}\right|_{chem} = \hat{R}_{m_i}(\overline{m}_1, ..., \overline{m}_N) + \hat{Q}_{m_i}(\overline{m}_i), \quad (37)$$

where $\overline{m}_i = \overline{q}_i (M_{air} / M_i)$ is used as the definition of the volumetric or molar mixing ratio, and $\hat{R}_{m_i} = R_{\varphi_i} / \overline{\rho}$ and $\hat{Q}_{m_i} = Q_{\varphi_i} / \overline{\rho}$ represent chemistry reactions and source terms in molar mixing ratio. In the remainder of this section, the over bars will be dropped for simplicity.

By using the kinetics laws for elementary reactions and by applying a mass balance to each species, the equation for the rate of change of each species concentration can be derived for a single cell:

$$\frac{dm_i}{dt} = P_i - L_i m_i \qquad i = 1, 2, ..., N \tag{38a}$$

with the initial conditions at $t = t_o$:

$$m_i(t_o) = m_{oi} \tag{38b}$$

where

$$P_i = \sum_{l=1}^{I_i} \upsilon_{i,l} r_l \tag{38c}$$

and

$$L_i m_i = \sum_{l=1}^{J_i} r_l \tag{38d}$$

In Eqs. (38a-d), P_i represents the production of species i, $L_i m_i$ represents the chemical loss of species i, $\upsilon_{i,l}$ is the stoichiometric coefficient for species i in reaction l, and r_l is the rate of chemical reaction l. The sum $l = 1...I_i$, in Eq. (38c) is over all reactions in which species i appears as a product, and in Eq. (38d) $l = 1...J_i$ is over all reactions in which species i appears as a reactant.

The rate of chemical reaction l can be expressed as the product of a rate constant k_l and a term that is a function of the concentration of the reactions. For elementary reactions, the concentration dependent term is the product of the reactant concentrations. In terms of mixing ratios, the rate of reaction l (r_l) takes one of the following forms:

$$r_l = \begin{cases} k_l m_i & \text{for first} - \text{order reactions} \\ k_l m_i m_j & \text{for second} - \text{order reactions} \\ k_l m_i m_j m_k & \text{for third} - \text{order reactions} \end{cases} \tag{39}$$

Equation (38a) forms a system of ordinary differential equations (ODEs) that must be solved numerically. Two sources of difficulty arise in obtaining solutions to it when applied to atmospheric chemistry problems. First, the system is nonlinear due to the contributions of the second- and third-order reactions. Second, the system of equations is "stiff" because of the widely varying time scales of the chemical reactions and the complex interactions among species. Mathematically, a stiff ODE system is described as one in which all the eigenvalues of the system (in this case, Eq. 38a) are negative, and the ratio of the absolute value of the largest-to-smallest real parts of the eigenvalues is much greater than one. For atmospheric chemistry problems, the ratio is often greater than 10^{10}, making the system very stiff (Gong and Cho [83]). Because the stiff ODEs must be solved over tens of thousands of cells repeatedly for air quality simulations, efficient numerical methods must be employed. The use of a standard explicit method is often precluded because relatively small time steps are required to maintain numerical stability and obtain accurate solutions. On the other hand, classical implicit methods that are both accurate and stable are not often applied because of the high computational costs. As described above, the CMAQ contains four different solvers – two generalized gas-phase solvers (SMVGEAR and QSSA) and two mechanism specific solvers (EBI and MEBI). Each solver offers varying levels of accuracy, computational efficiency, and flexibility. A brief description of each is included below, followed by a description of model simulation results that illustrate the relative differences in computational efficiency and numerical solution results.

Sparse Matrix Vectorized Gear (SMVGEAR) solver

Numerical solvers based on the algorithm developed by Gear [84] have traditionally been used to obtain accurate solutions to stiff ODE problems. The technique is an implicit method that incorporates an automatic time step size and error control and does not amplify errors from one step to another. SMVGEAR is a version of the Gear algorithm developed by Jacobson and Turco [85] to increase computational efficiency. SMVGEAR is most efficient when run on vector computers, but it is still faster than many other Gear codes on non-vector platforms because of the introduction of special sparse-matrix techniques. The SMVGEAR algorithm implemented in the CCTM is essentially the same as that developed by Jacobson and Turco [85], although fairly extensive changes were made to the original computer code to link the algorithm to the generalized chemical mechanism processor used in the CMAQ system and to make it conform to CMAQ coding conventions. As with most solvers, the accuracy of the

numerical solutions and the computational efficiency of SMVGEAR are affected by error tolerances that control solution accuracy. The SMVGEAR uses a relative tolerance to control the number of accurate digits and an absolute error tolerance to control noise level. In the CCTM implementation, a set of user-controllable tolerances are applied to all species, and default relative and absolute tolerances have been set to 10^{-3} and 10^{-9} ppm, respectively.

Quasi Steady-State Approximation (QSSA) solver

The QSSA solver is a low-order, explicit solver that exhibits good stability for stiff systems. Although less accurate than the Gear solver, many variants have been developed and used in 3-dimensional AQMs (e.g., [86,6,10,87]). The QSSA solver used in the CCTM is based on the algorithm in the Regional Oxidant Model [6,88]. The CMAQ QSSA solver is completely generalized and has been linked to the CMAQ generalized chemical mechanism processor. Thus, no *a priori* assumptions about reaction time scales are made, nor is any mechanism-specific species lumping performed or species steady-state relations assumed. As a consequence, other non-generalized QSSA methods that employ lumping techniques to ensure mass conservation may be somewhat faster and more accurate. Nevertheless, its generality facilitates the inclusion of multiple chemical mechanisms in the CMAQ modeling system. The accuracy of the solver can be controlled by, for example, the upper time step limit (τ_{ulim}). After a series of tests, we currently recommend τ_{ulim} =1 minute. Note that τ_{ulim} = 5 minutes was previously used as the default.

Euler Backward Iterative (EBI) and Modified Euler Backward Iterative (MEBI) solvers

The EBI solver developed by Hertel et al. [89] uses functional iteration to obtain a solution to the implicit Euler backward approximation. To speed up convergence, groups of species that are strongly coupled are isolated from the rest of the species in the mechanism and analytical expressions are derived to compute their concentrations. These results are then used with the backward Euler approximation to obtain estimates of the concentrations for the remainder of the mechanism species. The concentrations of the group species and the Euler backward species are updated using the results from the previous iteration until convergence is achieved. Hertel et al. [89] applied this technique to a particular version of the CB4 mechanism and found it to be more efficient than a version of the QSSA that employed species lumping. They point out that the procedure applies to the specific mechanism described, and may need to be modified when implemented for other mechanisms.

In CMAQ, two variants of the EBI approach are included. First, the same approach used by Hertel et al. has been applied to the CMAQ CB4 mechanism. Analytical expressions have been derived for four different groups of strongly coupled species: 1) NO, NO_2, O_3, and $O(^3P)$; 2) OH, HO_2, HONO, and HNO_4; and 3) NO_3 and N_2O_5, and 4) PAN and C_2O_3. The EBI iterative method is then used for the remaining species. The second variant (Modified EBI or MEBI) utilizes the method of Huang and Chang **[**90**]** to replace the analytical expressions for groups 1 and 2 with Newton-Raphson numerical solutions (analytical expressions are retained for groups 3 and 4). This approach is slower than the EBI approach because it requires matrix inversions, but it is somewhat more amenable to generalization. Even though the MEBI is slower than EBI, it is still significantly faster than either SMVGEAR or QSSA on scalar computers. To date, the MEBI approach has been developed for the following CMAQ mechanisms – CB4, SAPRC99, and the RADM2 variant utilizing the 4-product isoprene chemistry. The development of EBI solvers for other CMAQ mechanisms will be considered for future versions of CMAQ.

The accuracy of the solution obtained with EBI and MEBI is controlled by the size of the chemistry time step and convergence tolerances for each species for the Euler iterations. In addition, the MEBI solvers include convergence tolerances for the Newton-Raphson routine. In CMAQ, the time step has a default value of 2.5 minutes, and convergence tolerances are set for each species individually on the basis of the work done by Huang and Chang [90] and subsequent testing and comparison with results obtained using the SMVGEAR solver. These defaults were selected to provide relatively accurate solutions without greatly sacrificing computational efficiency.

8. AEROSOL AND AQUEOUS PROCESSES IN CMAQ

The original U.S. EPA PM NAAQS, issued in 1971 as total suspended particulates (TSP), were revised in 1987 to protect against human health effects associated with exposure to ambient PM having particle diameters less than 10 μm (PM_{10}). In 1997, a re-evaluation of the issue using newly available information provided key scientific bases for promulgation of additional standards for fine particles (i.e., those with diameters less than 2.5 μm, known as $PM_{2.5}$), while retaining current standards for PM_{10}. The vast majority of fine particulate mass across the Eastern U.S. and Canada is secondary in origin and sulfate is a dominant constituent involved with both the aerosol and aqueous

processes. In the following we describe CMAQ algorithms implemented for these processes.

8.1 Aerosol Processes in the CMAQ CTM

In the CCTM, aerosol particles are divided into two groups: fine particles and coarse particles. The fine particles resulting from combustion and secondary production processes are considered to have the same chemical composition and consist of sulfates, nitrates, organic carbon, and elemental carbon. The coarse particles consist of wind-blown dust (crustal materials), marine particles, and some anthropogenic contributions. Fine and course particles do not interact in this version of the aerosol module. We assume that the fine aerosol particles are in equilibrium with the ambient gas and vapor phase species at the ambient relative humidity and form an aqueous solution. Detailed descriptions of the aerosol algorithms and test results are available in Binkowski and Roselle [91] and, therefore, only a few highlights of the aerosol processes are described here.

The present implementation of the aerosol module in the CCTM is based on a modal aerosol modeling approach. Polydisperse fine particles are characterized by a bimodal lognormal distribution and coarse particles are described with a unimodal lognormal distribution. In the modal approach, particle distributions are characterized by the lognormal moments. The *k*-th moment of the distribution is defined as

$$M_k = \int_{-\infty}^{\infty} D^k n(\ln D) d(\ln D), \qquad (40)$$

where n is the number density and D is the particle diameter. M_0 is the total number (N) of particles, M_2 is proportional to the total particulate surface area, and M_3 is proportional to the total particulate volume. The size distribution of fine particles changes in response to coagulation between particles, growth by condensation from gas/vapor phase species, new particle production from vapor phase precursors, transport of particles, and emission of new particles. The algorithms describing these processes are extensions of those developed for the Regional Particulate Model (RPM) [92]. As in the RPM, $PM_{2.5}$ in the CCTM is represented by the two lognormal sub-distributions called the Aitken and accumulation modes. The total particle size distribution is represented by the superposition of Aitken and accumulation modes with variable standard deviations and a coarse mode distribution with a fixed standard deviation of 2.2. The ISORROPIA model is used to calculate the thermodynamic equilibrium between inorganic

aerosol species (nitrate, ammonium, and water) and gas-phase concentrations [93, 94].

To demonstrate the evolutionary characteristics of the distributions of moments, we reproduced Figure 4 from Figures 1, 2, and 3 in Binkowski and Roselle [91]. The initial values for the clear, urban and hazy cases are provided following Seigneur et al. [95]. For the clean background, the number density distribution has a typical peak in the size range 0.01 to 0.1 μm or in the Aitken mode (Whitby [96]). The volume (mass) distribution peaks in the size range larger than 1 μm, showing that the main contribution to the total mass of the total size distribution is in the coarse mode. The second peak for volume (mass) is in the size range 0.1 to 1.0 μm or in the accumulation mode. The size distribution of surface area has contributions from both the Aitken and accumulation modes. For a typical polluted urban case, the number distribution peaks in the Aitken mode. The mass, however, peaks in the accumulation mode with a second peak in the coarse mode. Surface area has a small contribution from the Aitken mode, but is dominated by particles in the accumulation mode. The initial number concentrations for the hazy case show a peak in the 0.03 μm size range. The curves shown as the final values in Figure 4 are from the single-cell calculation with the CMAQ aerosol module after 12-hours of evolution. Only the coagulation process was present for the clear and urban cases while the hazy case was with the added sulfate condensation process. This ideal simulation was designed to stress the aerosol dynamics algorithms. For the clear case, we see only little change in the number density distribution. No changes in the distributions of volume and surface area are noticeable. For the urban case, change in the number concentration is dramatic and the volume and surface area distributions show corresponding evolutionary patterns. The hazy case shows tremendous increase in the region below 0.1 μm (i.e., Aitken mode) for all three moments, representing the effects of the condensation process. The important point to be emphasized here is that condensation of secondary material upon an existing particle distribution can have a much larger effect on changing the distribution than coagulation.

The effects of aerosol chemistry and dynamics on aerosol species concentrations are solved with a fractional time step method:

$$\left.\frac{\partial \varphi_i^*}{\partial t}\right|_{aero} = \sqrt{\hat{\gamma}}\, R_{aero_i}(\overline{\varphi}_1,...,\overline{\varphi}_N) + \sqrt{\hat{\gamma}}\, Q_{aero_i} - \hat{v}_g \frac{\partial \varphi_i^*}{\partial \xi} \qquad (41)$$

where R_{aero_i} represents processes such as new particle formation and growth and depletion of existing particles. Q_{aero_i} stands for all the external sink and source terms, and $\hat{v}_g$ is the contravariant sedimentation velocity. The generic concentration units for the aerosol process are [μg m^{-3}] (density) for aerosol mass, [number m^{-3}] for aerosol particle number density, and [m^2 m^{-3}] for surface area density. Because the aerosol process is called between the pair of couple/decouple calls, the input concentration is already decoupled and the following set of governing equations are solved in the aerosol process module:

$$\left.\frac{\partial \varphi_i}{\partial t}\right|_{aero} = R_{aero_i}(\overline{\varphi}_1,...,\overline{\varphi}_N) + Q_{aero_i} - \hat{v}_g \frac{\partial \varphi_i}{\partial \xi} \qquad (42)$$

The key algorithms simulating aerosol processes in the CCTM include: (1) aerosol removal by size-dependent dry deposition; (2) aerosol-cloud droplet interaction and removal by precipitation; (3) new particle formation by binary homogeneous nucleation in a sulfuric acid/water vapor system; (4) the production of an organic aerosol component from gas-phase precursors; and (5) particle coagulation and condensational growth.

Two important aerosol processes included in the CCTM involve the heterogeneous conversion of N_2O_5 to HNO_3 and the condensation of semi-volatile compounds resulting from the atmospheric oxidation of some organic compounds. In CMAQ, the heterogeneous conversion of N_2O_5 to HNO_3 is based on the method of Dentener and Crutzen [97], in which the loss of gaseous N_2O_5 is treated as a pseudo first-order process. The rate coefficient for this process is expressed as a function of the gas-phase diffusion coefficient of N_2O_5, the radius of the aerosol particle, the aerosol surface area, and the reaction probability. Dentener and Crutzen used a value 0.1 for the reaction probability, but CMAQ incorporates the method used by Riemer et al. [98] in which the reaction probability is tied to the nitrate composition of the aerosol and varies between 0.02 and 0.002. It is based on experimental results suggesting that the heterogeneous N_2O_5 conversion rate is inhibited by the presence of nitrates in the aerosol phase [99]. At the end of the aerosol time step, the amount of HNO_3 retained in the aerosol phase is determined from thermodynamic equilibrium considerations. The portion of N_2O_5 that is not converted is assumed to remain in the gas-phase.

The formation of secondary aerosols in CMAQ occurs via the condensation of semi-volatile compounds that are produced from the gas-phase reactions of alkanes, aromatics, cresols, and terpenes. Because of their volatility,

the semi-volatiles are also capable of evaporating from the particle-phase back to the gas-phase when temperatures increase. The SOA gas-particle partitioning model used in CMAQ is based on the Secondary Organic Aerosol Model (SORGAM) described by Schell et al. [100]. However, the gas-phase chemistry yields and the saturation concentrations of the semi-volatiles that are used in CMAQ are derived from the following sources – Odum et al. [101] for aromatics, Griffin et al. [102] for terpenes, and Strader et al. [103] for alkanes and cresols. Finally, the CMAQ SOA algorithm has been implemented in such a manner as to distinguish between SOA produced from anthropogenic versus biogenic sources.

The aerosol species predicted in the CCTM are mass concentrations of Aitken and accumulation modes sulfate, ammonium, nitrate, anthropogenic secondary organic aerosol, biogenic secondary organic aerosol, elemental carbon aerosol, respectively. For the coarse mode, the CCTM provides concentrations of unspecified species (anthropogenic aerosol, marine aerosol, and soil derived aerosol mass). Unspeciated number concentrations and surface areas of the Aitken and accumulation modes, respectively, are predicted as well as the mass concentrations of water in Aitken and accumulation modes.

8.2 Cloud Module

The current subgrid cloud scheme in CMAQ was derived from the diagnostic cloud model in RADM version 2.6 (Dennis et al. [104], Walcek and Taylor [105], Chang et al. [10], [71]). The CMAQ cloud module includes parameterizations for subgrid convective precipitating and non-precipitating clouds and grid-scale resolved clouds. It includes an aqueous chemistry model for sulfur, and includes a simple mechanism for scavenging. The rate of change in pollutant concentrations due to cloud processes is given by

$$\left.\frac{\partial \overline{m}_i}{\partial t}\right|_{cld} = \left.\frac{\partial \overline{m}_i}{\partial t}\right|_{subcld} + \left.\frac{\partial \overline{m}_i}{\partial t}\right|_{rescld} \qquad (43)$$

where subscripts *cld*, *subcld*, and *rescld* represent cloud, sub-grid scale cloud, and resolved cloud, respectively. The sub-grid cloud effects are accounted for once an hour on the half hour while the resolved cloud effects are determined at each call. In CMAQ, the subgrid clouds are considered only for horizontal grid resolutions on the order of 8 km or more, and therefore for grid resolutions below 8 km, only resolved clouds are represented.

The effects of subgrid clouds on grid-averaged concentrations are modeled by the mixing, scavenging, aqueous chemistry, and wet deposition of a "representative cloud" within the grid cell. Subgrid clouds can be either precipitating or non-precipitating, and the non-precipitating sub-grid clouds are further categorized as pure fair weather clouds and non-precipitating clouds coexisting with precipitating clouds. RADM used the total precipitation (sum of convective and nonconvective precipitation amounts provided by MM5) to drive the sub-grid cloud model. The CMAQ implementation differs from RADM in that only the convective precipitation amounts from a meteorology model (in this case MM5) are used to drive the sub-grid precipitating cloud, and the nonconvective precipitation is used in the resolved cloud model. The mixing in the sub-grid convective cloud takes into account air transported vertically – from below the cloud, entrained from above the cloud (for precipitating clouds), and entrained from the sides of the cloud.

In the resolved cloud module, the pollutants, cloud, and rain are assumed to be uniformly distributed in a grid cell. No additional cloud dynamics are considered for the resolved cloud in CMAQ because any convection and/or mixing would have been treated by the vertical transport. The CMAQ resolved cloud model processes scavenging, aqueous chemistry, and wet deposition with the total condensed cloud and rain water using the same procedures as in the sub-grid clouds.

Scavenging and wet deposition

Pollutant scavenging is calculated by two methods, depending upon whether the pollutant participates in the cloud water chemistry. For those pollutants that are absorbed into the cloud water and participate in the cloud chemistry, the amount of scavenging depends on Henry's law constants, dissociation constants, and cloud water pH. For pollutants that do not participate in aqueous chemistry, the model uses the effective Henry's Law equilibrium equation to calculate ending concentrations and deposition amounts. The rate of change for in-cloud concentrations (m_i^{cld}) for each pollutant following the cloud time scale (τ_{cld}) is given by:

$$\left.\frac{\partial \overline{m_i}^{cld}}{\partial t}\right|_{scav} = \overline{m_i}^{cld}\left(\frac{e^{-\alpha_i \tau_{cld}} - 1}{\tau_{cld}}\right) \qquad (44)$$

where α_i is the scavenging coefficient for the pollutant. For sub-grid convective clouds, τ_{cld} is 1 hour and for grid

resolved clouds it is equal to the CMAQ's synchronization time step. For gases, the scavenging coefficient is given by:

$$\alpha_i = \left[\tau_{washout} \left(1 + \frac{TWF}{H_i^*} \right) \right]^{-1} \quad (45)$$

where H_i^* is the effective Henry's Law coefficient for the pollutant, *TWF* is the total water fraction given by:

$$TWF = \frac{\rho_{H_2O}}{\overline{W}_T RT} \quad (46)$$

where ρ_{H_2O} is the density of water, $\overline{W}_T$ is the mean total water content (kg m^{-3}), R is the Universal gas constant, and T is the in-cloud air temperature (K). The washout time, $\tau_{washout}$ represents the amount of time required to remove all of the water from the cloud volume at the specified precipitation rate (P_r), and is given for the cloud with thickness Δz_{cld}:

$$\tau_{washout} = \frac{\overline{W}_T \Delta z_{cld}}{\rho_{H_2O} P_r} \quad (47)$$

The accumulation mode and coarse mode aerosols are assumed to be completely absorbed by the cloud and rain water. Therefore, the scavenging coefficients for these two aerosol modes are simply a function of the washout time:

$$\alpha_i = \left[\tau_{washout} \right]^{-1} \quad (48)$$

The Aitken mode aerosols are treated as interstitial aerosol and are slowly absorbed into the cloud/rain water. An assumption used is that organics influence neither the water content nor the ionic strength of the system. Thus, in the current release of CMAQ, only the equilibrium of the sulfate, nitrate, ammonium, and water system is considered. The equilibrium and associated constants are based upon Kim et al. [106]. Refer to Binkowski and Roselle [91] and Binkowski [107] for the details on the aerosol-phase chemistry.

The wet deposition algorithms in CMAQ were taken from the RADM. In the current implementation, deposition is accumulated over 1-hour increments before being written to the output file. The wet deposition amount of chemical species i ($wdep_i$) depends upon the precipitation rate (P_r) and the cloud water concentration (m_i^{cld}):

$$wdep_i = \int_0^{\tau_{cld}} \overline{m_i}^{cld} P_r dt \quad (49)$$

8.3 Aqueous-Phase Chemistry

The aqueous chemistry model evolved from the original RADM (Chang et al. [10]). The model considers the absorption of chemical compounds into the cloud water; the amount that gas-phase species absorb into the cloud water depends on thermodynamic equilibrium, while accumulation-mode aerosols are considered to have been the nucleation particles for cloud droplet formation and are completely absorbed into the cloud water. Then the model calculates the dissociation of compounds into ions, oxidation of S(IV) to S(VI), and wet deposition. The species that participate in the aqueous chemistry are given in Table 3. This version of the aqueous chemistry model differs from the Walcek and Taylor [105] scheme in that it tracks contributions from gases and aerosols separately. It also considers the scavenging of interstitial aerosols, and it allows for variable-length cloud time scales.

9. PLUME-IN-GRID PROCESS

Substantial anthropogenic emissions of nitrogen oxides (NO_x) and/or sulfur oxides (SO_x) are released from individual point sources into plumes whose horizontal dimension remains considerably smaller than the typical size of regional photochemical model grid cells (e.g. 20-40 km) for some time after release. The plume-in-grid (PinG) approach in CMAQ, which is based on a Lagrangian reference frame, was specifically designed to address the need to more realistically resolve the spatial scale of plumes emanating from isolated, major elevated point source emitters (MEPSEs) within an Eulerian regional grid framework. Since the traditional Eulerian grid modeling approach has been to instantly mix point source emissions into an entire grid cell volume, it bypasses the diffusion-limited, chemical evolution phase occurring in sub-grid scale plumes during their transport downwind, which may strongly impact the photochemically-formed species concentrations in regional model grid cells. The chemical mixture in fresh, point source plumes, particularly of fossil-fuel power plants, can be characterized to be in a high NO_x / low VOC regime, while the ambient environment surrounding a plume is often in the opposite chemical regime. The over dilution and mixing of high NO_x point source emissions into large grid cells can prematurely initiate rapid ozone production. However, the PinG technique overcomes this undesirable feature since it provides spatial resolution within a plume and also simulates the gradual horizontal growth experienced by real-world plumes.

The detailed mathematical formulation of the processes treated in the PinG approach is documented in Gillani and Godowitch [108]. The CMAQ/PinG processes were implemented in the two key modeling components; the plume dynamics model (PDM) processor and the PinG module, which is embedded within the CCTM. The PDM processor simulates plume rise, and generates the vertical/horizontal dimensions and grid positions of a series of moving plume sections released at hourly intervals from the location of each MEPSE point source. The PDM data file is used by the Lagrangian PinG module, which is fully coupled with the Eulerian CCTM and is executed concurrently with the grid model (Godowitch and Young [109]). The PinG module treats the horizontal/vertical dispersion of plume concentrations, entrainment of background concentrations and horizontal eddy diffusion (*disp*), emissions into the plume to account for the plume chemistry effects of anthropogenic and biogenic surface emissions (*emis*), plume chemistry to appropriately simulate the chemical evolution within the plume (*chem*), and surface dry deposition to account for the removal of pollutant within the plume (*dep*) which can be expressed by

$$\frac{\partial m_p}{\partial t} = \left.\frac{\partial m_p}{\partial t}\right|_{disp} + \left.\frac{\partial m_p}{\partial t}\right|_{emis} + \left.\frac{\partial m_p}{\partial t}\right|_{chem} + \left.\frac{\partial m_p}{\partial t}\right|_{dep}, \qquad (50)$$

where m_p is the concentration of the sub-grid plume (in molar mixing ratio). Horizontal resolution across a plume section is achieved with a contiguous array of attached plume cells. Currently, each plume section is composed of 10 plume cells in addition to a left and a right boundary cell. The plume boundary conditions, representing the ambient background, are provided throughout the simulation by the appropriate CCTM grid concentrations. Currently in PinG, the plume cells represent a single vertical layer. Each plume cell in a particular plume section has the same plume bottom and a common top height. When the sub-grid scale phase of the plume simulation has been completed, the PinG module updates grid scale concentrations according to

$$\left.\frac{\partial m_i}{\partial t}\right|_{ping} = \frac{\delta V_p}{\delta V}\frac{\partial (m_p - m_{ibg})}{\partial t} \qquad (51)$$

where m_{ibg} is the background concentration and δV_p is the volume of plume in a grid cell with volume δV. Currently, gaseous species are treated with the PinG module; however, an aerosol/particulate module is now being incorporated. For consistency with the grid model, PinG applies the same gas-phase chemical mechanisms and chemical solvers used by the CCTM.

The CCTM/PinG model was applied on a domain encompassing the greater Nashville, Tennessee region. Model simulations were performed for selected case study days from July 1995, which coincided with the Southern Oxidant Study (SOS) field study program being conducted in the greater Nashville area. In particular, five major point sources exhibiting a range of NO_x emission rates were selected for the PinG treatment. Figure 5 compares PinG modeled species concentrations to observed plume concentrations obtained from a horizontal aircraft traverse intercepting multiple point source plumes. These example results are encouraging as PinG concentrations for O_3, NO_y, and SO_2 and others presented in Godowitch [110] are in reasonable agreement with plume measurements for three different point source plumes intercepted at different downwind distances along the same flight track. Further quantitative analyses from several case study days are underway in order to provide an initial assessment of the ability of PinG to replicate the photochemical behavior of ozone and other pollutant species in point source plumes.

10. APPLICATION EXAMPLES OF CMAQ CTM

The current version of CMAQ is capable of studying a variety of air quality problems such as tropospheric ozone, fine particles, acid deposition, nutrients, and visibility degradation for urban to regional scales. Here we present a few application examples of CMAQ simulations for ozone and PM air quality problems on urban/regional scales.

10.1 Applications to continental United States

With the recent changes in U.S. National Ambient Air Quality Standards (NAAQS) to include daily and annual fine particulate matter (PM2.5) and 8-hr averaged ozone, there is growing interest in the ability of air quality models to simulate these pollutants across the U.S. The most recent version (September 2003) of the CMAQ model was used to perform continental U.S. (CONUS) scale applications for two different time periods: 15 June – 16 July 1999 and 4 January – 19 February 2002. A winter period is included here since PM2.5 is a year-round air quality issue for many regions of the U.S., although ozone is principally a warm-season problem.

Meteorological data were prepared for both seasonal applications with the MM5 model [12] and were processed by the MCIPv2.2 program for CMAQ model use. The July 1999 application used 32-km horizontal grid sizes and 30 vertical model layers extending to 10^4 Pascal. The winter 2002 application used 36-km horizontal grid sizes and 34

vertical model layers. The corresponding CMAQ model grid used the same horizontal grid resolutions as MM5. MM5's meteorological data were converted to a 21-layer vertical structure by the MCIP for CMAQ for the July 1999 application and 24 layers for CMAQ's January 2002 application.

Source emissions data were obtained from the 1999 EPA National Emissions Inventory (NEI) [111] and processed through the SMOKE emissions modeling system [24, 25] for CMAQ's use. Hourly gridded emissions of gas-phase SO_2, CO, NO, NO_2, NH_3, and VOC were included, as were anthropogenic PM2.5 emissions subdivided into particulate SO_4, NO_3, organic carbon (OC), and elemental carbon (EC). Mobile source emissions were processed using the MOBILE5b model and biogenic source emissions were estimated from the BEISv3.11; both models are incorporated in the SMOKE system. VOC emissions were split into the appropriate organic categories for the SAPRC-99 chemical mechanism used in these CMAQ model applications. Initial and boundary condition chemical data were set at relatively clean tropospheric background levels. The CMAQ model was run to produce hourly estimates of gas- and particle-phase chemical trace species for the four-week summer period and six-week winter period.

Data used for model evaluation were obtained from EPA's AIRS database (hourly ozone - >700 stations used, mostly urban/suburban), from the IMPROVE (Interagency Monitoring of Protected Visual Environments) network for daily averages (every third day) of SO_2, NO_3, PM2.5, OC, and EC from 50 rural sites, from the STN (Speciated Trends Network) for 2002 only for daily averages (every third day) of SO_2, NO_3, PM2.5, NH_4, OC, and EC from 60 urban sites, and from the SEARCH (Southeast Aerosol Research and Characterization) research sites for 1999 only for daily and hourly averages of PM2.5 and its components.

CMAQ model results for maximum daily 8-hr ozone concentrations are shown in Figure 6 and Table 4 for the 4-week summer 1999 application period. There were over 23,000 data points in the analysis covering over 700 monitoring stations. While the scatterplot in Figure 6 illustrates considerable variability in the model/observations intercomparison, correlations are strong (r=0.75) and nearly all comparisons are within a factor of two. CMAQ model's mean bias is 4.3 ppb (or 8.7% as a normalized mean bias). The frequency distribution in Table 4 indicates that the model tends to slightly underpredict at the highest ambient ozone levels and overpredict at the low end of the distribution.

Model results for daily average PM2.5 concentrations for the 4-week summer period are shown in Figure 7 and Table 5. CMAQ model results from PM2.5 are obtained by summing the simulated aerosol concentrations of sulfates, nitrates, ammonium, organic and elemental carbon, and uncategorized aerosol. Results shown on the scatterplot in Figure 7 indicate that model and observed aerosol concentrations for the summer period are considerably higher in the eastern sites than in the west. Most data points fall within the 2:1 bounds and are correlated (r=0.71), with a small mean bias (-0.73 $\mu g\ m^{-3}$, or -9.8% as a normalized mean bias). Among the principal aerosol components (not shown), sulfate tends to be slightly overpredicted and organic aerosols are slightly underpredicted during this time period.

Similar results for the 6-week wintertime period are shown in Figure 8 and Table 6. More observations are included in this analysis because of the longer simulation period, and also the STN network became operational in the urban areas. The scatterplot shows that the urban STN data are generally higher than the rural IMPROVE data, and that some of the western sites have considerably higher aerosol concentrations than eastern sites (mostly because of high aerosol nitrate). The CMAQ model results generally were within a factor of two of observations with reasonable correlation (r=0.68) for the IMPROVE sites, but degraded correlation (r=0.37) for the STN sites. Mean biases were 1.49 $\mu g\ m^{-3}$ (40.3% as normalized bias) for the IMPROVE stations, and 0.51 $\mu g\ m^{-3}$ (4.1% as normalized bias) for the STN stations. The model has a tendency to overpredict concentrations in the winter application, although with differing results among the aerosol components. Nitrate aerosols were generally overpredicted, while organic aerosols were underpredicted. Some of the STN stations in the western U.S. showed much higher nitrate and organic aerosol concentrations than the CMAQ model predicted.

This application illustrates that the CMAQ model is able to simulate ozone and PM2.5 concentrations reasonably well over the CONUS. There are many challenges, especially in the area of aerosol modeling and interpreting the observational data for model evaluation. We have seen that different networks (urban vs. rural; east vs. west) can show quite different characteristics in the ambient aerosols. In the coming year, the CMAQ model will be used to simulate ozone and PM2.5 for the full year of 2001 over the CONUS for a more extensive model evaluation application.

10.2 Application to Houston-Galveston airshed high ozone events

Petroleum refinery emissions cause serious air quality problems as they are respiratory irritants causing significant health effects and suspected of elevating cancer risks in certain populations. Nitrogen oxides and various VOCs emitted by the petroleum refineries and associated chemical facilities are precursors for ozone and particulate pollutants. The Houston-Galveston Airshed (HGA) has become one of the most severe ozone non-attainment regions in the U.S. for both 1-hour ozone and the number of days above the National Ambient Air Quality Standard. The HGA contains about 50% of the nation's petrochemical capacity: 27.6 billion pounds of ethene (i.e., 52% of the nation's capacity) and 10.9 billion pounds of polymer grade propene (63% of the nation's capacity). These olefin emissions can affect ozone concentrations significantly. A variety of measurements made during the Texas Air Quality Study (TexAQS2000), August 15-September 19, 2000, [112] have detected massive and frequent spikes of ozone, which appear to be associated with releases of reactive unsaturated hydrocarbons (olefins) from the petrochemical industries. Compared with typical ozone evolution patterns in other U.S. cities, several monitoring sites in the area show distinct rapid transient high ozone events (THOEs) [113].

Meteorology and air quality of the episode

In summer, a high-pressure system persists over the Gulf of Mexico, leading to stagnant conditions in southeast Texas. Under this weak synoptic scale forcing, a land-sea breeze circulation becomes apparent and dominates the weather pattern in the Houston-Galveston area. The TexAQS 2000 episode had numerous exceedences (daily maximum ozone concentration greater than 125 ppb) in both Houston and Beaumont. There were six exceedence days in the HGA during the eight-day period, including a period of low ozone in the middle. The episode included several days with continued veering wind vectors causing flow reversal from morning to afternoon hours. These high ozone concentration events were frequently linked with a land-sea breeze circulation in the HGA.

August 25 showed light easterly winds resulting in maximum ozone at Crawford at the center of the Houston area. August 26, 27 and 28 were low ozone days. Stronger southeasterly sea breeze winds resulted in substantially lower ozone in the HGA and transported the diluted urban plume to Conroe. August 29, 30, and 31 showed light westerly winds followed by an afternoon sea breeze which positioned the ozone pool on the east side of the city at Mt. Belview, La Porte and Deer Park. September 1 had a relatively persistent westerly wind, which carried the

maximum ozone to the Baytown monitor and areas further east.

Model set up

We used MM5 simulations for the TexAQS 2000 period, provided by Nielsen-Gammon [114]. The simulation used a slab soil model with temporally varying soil moisture that was modified in the urban area to make it wetter and the rural areas to make it drier. MM5 Version3 Release 6 (MM5v3.6) was used. MM5 physics options applied include: 1) Grell cumulus scheme on the 108-, 36- and 12-km domains with 43-modeling layers, 2) MRF PBL scheme, 3) Dudhia simple ice microphysical scheme and cloud-radiation scheme. The first guess initialization field and boundary conditions were from the NCEP Eta model on the Eta212 (AWIP 40 km domain) and upper air analysis nudging was used. The wind profiler measurements during the TexAQS 2000 study period provided a unique opportunity to compare the MM5 simulation of planetary boundary layer (PBL) heights with the observations. It showed that MM5 simulated 10-30% higher PBL heights than observations. For air quality simulations, the 43-layer data have been reduced to 22-layer, where collapsing was applied for layers above 3-km height.

VOC emissions uncertainty in the Houston-Galveston area is large. It is suggested that alkene emissions in inventories are significantly (factor of 3-10) lower than those inferred from the aircraft measurements of formaldehyde concentration during the TexAQS 2000. With such emissions uncertainties, models will not accurately simulate observations. Here we tested the emissions inventory prepared by the Texas Commission for Environmental Quality (TCEQ), who modified the olefin emissions from several large point sources along the Houston Ship Channel to have similar magnitudes as the NO_x emissions.

CMAQ simulations including the CB-4 chemical mechanism were performed for the episode August 23 – September 1, 2000 at 36-, 12-, and 4-km resolutions. Boundary condition profiles were provided for 36-km runs, while boundary conditions for the nested 12-km and 4-km runs were provided from modeled results for the coarser domains. In order to provide more realistic initial conditions for the CMAQ runs, the model was allowed a spin-up time of two days, then, restarted with initial conditions from hour 24 of August 24, 2000. This procedure increased the reliability of the modeled results for August 23-24, 2000.

Simulated ozone concentrations

Figure 9 provides daily maximum ozone concentration fields from the simulation. The center of Houston and areas downwind of the Ship Channel industrial areas show high ozone concentrations. The highest value was 124 ppb, which occurred around hour 20:00 UTC on August 31, 2000. From the scatter diagram (Figure 10) comparing model simulation with the continuous air monitoring site (CAMS) data, we can identify the sites showing large discrepancies. They are the industrial Houston Regional Monitoring Sites 7, 8 and 11 (EPA482010807, EPA482010808, and EPA48710901, respectively) and the nearby Crawford (EPA482011037) and Deer Park (EPA482011039) sites. Time series plots for other sites show respectable simulation results for most of the simulation days (see Figure 11) except for the missed peak values on August 25 and 30. A subsequent analysis showed that the MM5 wind direction was off for several hours on August 25 and MM5-predicted PBL heights were about 30% too high on August 30 compared with the wind profiler measurements. The persistent differences between the simulated and observed concentrations at industrial and nearby sites demonstrate a need for further study in the emissions uncertainty in the HGA emissions inventories.

11. CONCLUDING REMARKS

The CMAQ model offers fully-functional multiscale and multi-pollutant air quality modeling capability by combining several distinct modeling techniques. We recast the governing atmospheric diffusion equation with a generalized coordinate system to accommodate many different meteorological coordinates and assumptions, including a fully compressible atmosphere. The CMAQ model can be configured to match the characteristics of the driving meteorological models. For example, hydrostatic meteorological models can be used to provide atmospheric characterization for regional scale simulations, and non-hydrostatic models can be used to simulate urban scale air quality, especially when there are significant topographic features. CMAQ uses a nesting capability, in addition to the plume-in-grid technique, to refine grid resolutions, passing the coarse-grid information to the nested fine-grid to fit the target air quality problem. Generalized gas-phase chemistry solvers, a linked aqueous reaction processor, and an aerosol module provide the multi-pollutant capability. The modular design allows exchange of science process modules and experiments with different model configurations.

The CMAQ model system aims to realize the “one atmosphere” approach to investigate various air quality issues for regional and urban scales, including ozone, particulate matter, acid deposition, and visibility. The capabilities for simulating air toxics, such as mercury,

dioxins, and other gas-, aerosol-, and metal-phase air toxics are now emerging in research versions of the CMAQ model. In addition to the examples of model application and evaluation discussed in this paper, other extensive CMAQ model evaluation studies are underway or planned for the future. We intend for a wider air quality modeling community to assist us in providing new state-of-science process modules for the CMAQ system to allow it to evolve for current and future air quality modeling assessment needs.

ACKNOWLEDGMENTS

The authors express their appreciation to many members of the air quality science community who contributed to the CMAQ model development project and to this paper. In particular we express great thanks to William Benjey, Jason Ching, Robin Dennis, Brian Eder, Gerald Gipson, James Godowitch, Joan Novak, Tanya Otte, Tom Pierce, Jonathan Pleim, Shawn Roselle, Jeffrey Young, Carey Jang and Mark Houyoux of U.S. EPA, Francis Binkowski, Adel Hanna, Kiran Alapaty, Carlie Coats, John McHenry, Rohit Mathur, Uma Shankar, and Aijun Xiu of Carolina Environmental Programs/UNC-Chapel Hill, Sharon LeDuc of NOAA, Noor Gillani and Arastoo Biazar of University of Alabama-Huntsville, Ruen Tang of CSC Corp., David Wong of SAIC Corp., Talat Odman of Georgia Institute of Technology, and Avraham Lacser, Sang-Mi Lee, and Seiji Sugata who were visiting scientists in the CMAQ development group at U.S. EPA. Thanks are also extended to EPA internal reviewers Deborah Luecken and Patrick Dolwick, and to Michelle Mebust who provided valuable editorial advice and assistance on the manuscript.

The United States Environmental Protection Agency, through its Office of Research and Development, funded and managed the research described here. This research has been subjected to Agency review and approved for publication.

REFERENCES

[1] Novak JH, Dennis RL, Byun DW, Pleim JE, Galluppi KJ, Coats CJ, Chall S, and Vouk MA (1995), EPA third-generation air quality modeling system, Volume 1: Concept, EPA 600/R95/084, U.S. Environmental Protection Agency, Research Triangle Park, NC.

[2] Dennis RL, Byun DW, Novak JH, Galluppi KJ, Coats CJ, and Vouk MA (1996), The next generation of integrated air quality modeling: EPA's Models-3, *Atmos. Environ.* **30**, 1925-1938.

[3] Byun DW, Young J, Gipson J, Godowitch J, Binkowski F, Roselle S, Benjey B, Pleim J, Ching J, Novak J, Coats C, Odman T, Hanna A, Alapaty K, Mathur R, McHenry J, Shankar U, Fine S, Xiu A, and Jang C (1998), Description of the Models-3 Community Multiscale Air Quality (CMAQ) model, Proceedings of the American Meteorological Society 78th Annual Meeting, Phoenix, AZ, 264-268.

[4] SAI (1990), User's Guide for the Urban Airshed Model. Volume I-V, Prepared by Systems Applications International, San Rafael, CA, Report No. SYSAPP-90/18a-e.

[5] McRae GJ, Russell AG, and Harley RA (1992), CIT Photochemical Airshed Model--Systems Manual, Carnegie Mellon University Report, Pittsburgh, PA.

[6] Lamb RG, (1983), A regional scale (1000 km) model of photochemical air pollution, Part 1: Theoretical formulation. EPA-600/3-83-035, U.S. Environmental Protection Agency, Research Triangle Park, NC.

[7] Carmichael GR, Peters LK, and Saylor RD (1991), The STEM-II regional-scale acid deposition and photochemical oxidant model: I. An overview of model development and applications, *Atmos. Environ.* **25A**, 2077-2090.

[8] Kumar N, Russell A, Segall E, and Steenkiste P (1996), Parallel and distributed application of an urban and regional multiscale model, *Computers and Chemical Engineering* **21**, 399–408.

[9] Environ (1998), User's Guide for Comprehensive Air Quality Model with Extensions (CAMx) Version 2.0. Environ Corporation, Novato, California.

[10] Chang JS, Brost RA, Isaksen ISA, Madronich S, Middleton P, Stockwell WR, and Walcek CJ (1987), A three-dimensional Eulerian acid deposition model: Physical concepts and formulation, *J. Geophy. Res.* **92,** 14681-14700.

[11] Anthes RA, and Warner TT (1978), Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev.* **106**, 1045–1078.

[12] Grell GA, Dudhia J, and Stauffer DR (1995), A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), NCAR Technical Note, NCAR/TN-398+STR, Boulder, CO, 138 pp.

[13] Chang JS, Jin S, Li Y, Beauharnois M, Chang K-H, Huang H-C, Lu C-H, Wojcik G, Tanrikulu S, and DaMassa J (1996), The SARMAP Air Quality Model. Part 1 of SAQM Final Report. California Air Resources Board, Sacramento, CA.

[14] Hass H, (1991), Description of the EURAD chemistry transport module (CTM) version 2. In Ebel A, Neubauer FM and Speth P (eds.), Report 83, Institute of Geophysics and Meteorology, Universiry of Cologne, Cologne, Germany.

[15] Grell GA, Emeis S, Stockwell WR, Schoenemeyer T, Forkel R, Michalakes J, Knoche R, and Seidl W (2000), Application of a multiscale, coupled MM5/chemistry model to the complex terrain of the VOTALP valley campaign, *Atmos. Environ.* **34**, 1435-1453.

[16] Jacobson MZ, Lu R, Turco RP, and Toon OB (1996), Development and application of a new air pollution modeling system – Part I: Gas-phase simulations, *Atmos. Environ.* **30**, 1939-1963.

[17] Vautard R, Beekmann M, Roux J, and Gombert D (2000), Validation of a deterministic forecasting system for the ozone concentrations over the Paris area, *Atmos. Environ*. **35**, 2449-2461.

[18] Robertson L, Langner J, och Engardt M (1999), An Eulerian limited-area atmospheric transport model, *J. Appl. Met.* **38**, 190-210.

[19] Källén E. (1996), HIRLAM documentation manual – System 2.5, 106pp. [Available from SMHI, SE-601 76 Norrköping, Sweden.]

[20] Brandt J, Christensen JH, Frohn LM, and Zlatev Z (2001), Operational air pollution forecast modelling by using the THOR system, *Physics and Chemistry of the Earth (B)* **26**, 117-122.

[21] Zlatev Z, Dimov I, and Georgiev K (1996), Three-dimensional version of the Danish Eulerian Model, *Zeitschrift für Angewandte Mathematik und Mechanik* **76**, 473-476.

[22] Fast JD, Zaveri RA, Xindi B, Chapman EG and Easter RC (2002), Effect of regional-scale transport on oxidants in the vicinity of Philadelphia during the 1999 NE-OPS filed campaign, *J. Geophy. Res.* **107**, 10.1029/2001JD000980.

[23] Pielke RA, Cotton WR, Walko RL, Tremback CJ, Lyons WA, Grasso LD, Nicholls ME, Moran MD, Wesely DA, Lee TJ, and Copeland JH (1992), A comprehensive meteorological modeling system–RAMS, *Meteorol. and Atmos. Physics*, **49**, 69-91.

[24] Coats CJ, Jr. and Houyoux MR (1996), Fast emissions modeling with the Sparse Matrix Operator Kernel Emissions Modeling System, Presented at The Emissions Inventory: Key to Planning, Permits, Compliance, and Reporting, Air & Waste Management Association, New Orleans, LA, September 1996.

[25] Houyoux MR and Vukovich JM (1999), Updates to the Sparse Matrix Operator Kernel Emission (SMOKE) modeling system and integration with Models-3, presented at the Emission Inventory: Regional Strategies for the Future, October 26-28, Raleigh, NC, Air and Waste Management Association.

[26] Coats CJ, Trayanov A, McHenry JN, Xiu A, Gibbs-Lario A, and Peters-Lidard CD (1999), An extension of the EDSS/Models-3 I/O API for coupling concurrent environmental models, In Applications to Air Quality and Hydrology, Preprints, 15th IIPS Conference, Amer. Meteor. Soc., Dallas, TX., January 10-15, 1999.

[27] Rew RK and Davis GP (1990), NetCDF: an interface for science data access, *IEEE Computer Graphics and Applications* **10**, 76-82.

[28] Byun DW (1999), Dynamically consistent formulations in meteorological and air quality models for multi-scale atmospheric applications: Part I. Governing equations in generalized coordinate system, *J. Atmos. Sci.* **56**, 3789-3807.

[29] Byun DW (1999), Dynamically consistent formulations in meteorological and air quality models for multi-scale atmospheric applications: Part II. Mass conservation issues, *J. Atmos. Sci.* **56**, 3808-3820.

[30] Ooyama KV (1990), A thermodynamic foundation for modeling the moist atmosphere, *J. Atmos. Sci*. **47**, 2580-2593.

[31] Seaman NL (2000), Meteorological modeling for air-quality assessments, *Atmos. Environ*. **34**, 2231-2259.

[32] Stauffer DR and Seaman NL (1994) Multiscale four-dimensional data assimilation, *J. Appl. Meteor.*, **33**, 416-434.

[33] Tremback CJ (1990), Numerical simulation of a mesoscale convective complex: model development and numerical results, Ph.D. dissertation, Colorado State University.

[34] Sugata S, Byun DW, and Uno I (2000), Simulation of sulfate aerosol in East Asia using Models-3/CMAQ with RAMS meteorological data, in Millennium NATO/CCMS International Technical Meeting on Air Pollution Modelling and Its Applications, Boulder, CO, May 15-19.

[35] Seaman NL (1995), Status of meteorological pre-processors for air-quality modeling, In proceedings of the *International Conf. on Particulate Matter*, Pittsburgh, PA, Air & Waste Management Association, 639-650.

[36] Vogel B, Fiedler F, and Vogel H (1995), Influence of topography and biogenic volatile organic compounds emission in the state of Baden-Wurttemberg on ozone concentrations during episodes of high air temperatures, *J. Geophys. Res*. **100**, 22,907-22,928.

[37] Xiu A, Mathur R, Coats C, and Alapaty K (1998), On the development of an air quality modeling system with integrated meteorology, chemistry, and emissions, in *Proceedings of the International Symposium on Measurement of Toxic and Related Air Pollutants*, Research Triangle Park, North Carolina, 144-152.

[38] Dudhia J, Gill D, Klemp J, and Skamarock W (1998), WRF: Current status of model development and plans for the future, in Preprints of the Eighth PSU/NCAR Mesoscale Model User's Workshop, Boulder, Colorado, 15-16 June.

[39] Klemp JB, Skamarock WC, and Dudhia J (2001), Conservative split-explicit time integration methods for the compressible nonhydrostatic equations, available at http://www.mmm.ucar.edu/individual/skamarock/wrf_equations_eulerian.pdf.

[40] Benjey WG, and Moghari NM (1996), Functionality of an integrated emission preprocessing system for air quality modeling: the Models-3 emission processor, in *The Emissions Inventory: Programs & Progress*. VIP-56, Proceedings of a Specialty Conference, Research Triangle Park, NC, October 11-13, 1995. U.S. EPA, Research Triangle Park, NC, and A&WMA, Pittsburgh, 463-471.

[41] Benjey WG, Godowitch JM, and Gipson GL (1999), Emission subsystem, Chapter 4 in Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, edited by D. Byun and J.K.S. Ching, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[42] Fratt DB, Mudgett DF, and Walters RA (1990), The 1985 NAPAP emissions inventory: Development of temporal allocation factors, EPA-600/7-89-010d, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., 209 pp.

[43] Moody T, Winkler JD, Wilson T, and Kersteter S (1995), The development and improvement of temporal allocation factor files, EPA-600/R-95-004, U.S. Environmental Protection Agency, Office of Research and Development.

[44] Carter WPL (2000), Implementation of the SAPRC-99 chemical mechanism into the Models-3 framework, Final report to the U.S. EPA.

[45] Pierce T, Geron C, Bender L, Dennis R, Tennyson G, and Guenther A (1998), The influence of increased isoprene emissions on regional ozone modeling, *J. Geophys. Res.*, 25611-25629.

[46] Guenther, A, Geron C, Pierce T, Lamb B, Harley P, and Fall R (2000), Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America, *Atmos. Environ*., **34**, 2205-2230.

[47] U.S. EPA (2002), User's guide to Mobile6.1 and Mobile 6.2: Mobile source emission factor model, EPA 420-R-02-028, Office of Air And Radiation, Office of Transportation and Air Quality, 264 pp.

[48] Toon OB, Turco RP, Westphal D, Malone R, and Liu MS (1988), A multidimensional model for aerosols: Description of computational analogs, *J. Atmos. Sci*., **45**, 2123-2143.

[49] Pleim JE (1990), Development and application of new modeling techniques for mesoscale atmospheric chemistry, Ph.D. Thesis, State University of New York at Albany, Albany, New York.

[50] Jeffries HE and Tonnesen S (1994), A comparison of two photochemical reaction mechanisms using mass balance and process analysis, *Atmos. Environ.*, **28**(18), 2991-3003.

[51] Jang JC, Jeffries HE, Byun DW, and Pleim JE (1995), Sensitivity of ozone to model grid resolution–I. Application of high-resolution Regional Acid Deposition Model, *Atmos. Environ.*, **29**(21), 3085-3100.

[52] Jang JC, Jeffries HE, and Tonnesen S (1995), Sensitivity of ozone to model grid resolution–II. Detailed process analysis for ozone chemistry, *Atmos. Environ.*, **29**(21), 3101-3114.

[53] Colella P and Woodward PR (1984), The piecewise parabolic method (PPM) for gas-dynamical simulations, *J. Comp. Phys*., **54**, 174-201.

[54] Bott A (1989), A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, *Mon. Wea. Rev.* **117**, 1006-1015.

[55] Odman MT (1998), Research on numerical transport algorithms for air quality simulation models, EPA Report. EPA/660/R-97/142, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[56] Byun DW, Young J, Pleim J, Odman MT, and Alapaty K (1999), Numerical transport algorithms for the community multiscale air quality (CMAQ) chemical transport model in generalized coordinates, Chapter 7 in Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, edited by Byun DW and Ching JKS, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[57] Smagorinsky J (1963), General circulation experiments with the primitive equations: 1. The basic experiment, *Mon. Wea. Rev.* **91**, 99-164.

[58] Stull RB (1988), *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, 666 pp.

[59] von Rosenberg D (1969), *Methods for the Numerical Solution of Partial Differential Equations*, Elsevier Publishing Co., p. 113

[60] Pleim JE and Chang J (1992), A non-local closure model for vertical mixing in the convective boundary layer, *Atmos. Env.,* **26A**, 965-981.

[61] Chang JS, Jin S, Li Y, Beauharnois M, Lu CH, Huang HC, Tanrikulu S, and DaMassa J (1997), *The SARMAP air quality model.* Final Report, SJVAQS/AUSPEX Regional Modeling Adaptation Project, 53 pp. [Available from California Air Resources Board, 2020 L Street, Sacramento, California 95814].

[62] Brost RA and Wyngaard JC (1978), A model study of the stably stratified planetary boundary layer, *J. Atmos. Sci.,* **35**, 1427-1440.

[63] Hass H, Jakobs HJ, Memmescheimer M, Ebel A, and Chang JS (1991), Simulation of a wet deposition case in Europe using the European Acid Deposition Model (EURAD), in *Air Pollution Modelling and Its Applications*, Vol. VIII (edited by van Dop H and Steyn DG), Plenum Press, 205-213.

[64] Byun DW and Dennis RL (1995), Design artifacts in Eulerian air quality models: Evaluation of the effects of layer thickness and vertical profile correction on surface ozone concentrations, *Atmos. Environ*., **29**, 105-126.

[65] Wesely ML (1989), Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmos.Environ.*, **23**, 1293-1304.

[66] Pleim JE, Clarke JF, Finkelstein PL, Cooter EJ, Ellestad TG, Xiu A, and Angevine WM (1996), Comparison of measured and modeled surface fluxes of heat, moisture and chemical dry deposition, in *Air Pollution Modeling and its Applications XI*, edited by Gryning and Schiermeier, Plenum Press, New York.

[67] Byun D, Pleim J, Tang R, and Bourgeois A (1999), Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system, edited by D Byun and JKS Ching, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[68] Madronich S (1987), Intercomparison of NO2 photodissociation and UV radiometer measurements, *Atmos. Environ*., **21**, 569-578.

[69] Joseph JH, Wiscombe WJ, and Weinman JA (1976), The delta-Eddington approximation for radiative flux transfer, *J. Atmos. Sci.*, **33**, 2452-2459.

[70] World Meteorological Organization (1986), Atmospheric ozone 1985: Assessment of our understanding of the processes controlling its present distribution and change, WMO Rep. No. 16; Global Ozone Research and Monitoring Project, Geneva, Switzerland.

[71] Chang JS, Binkowski FS, Seaman NL, Byun DW, McHenry JN, Samson PJ, Stockwell WR, Walcek CJ, Madronich S, Middleton PB, Pleim JE, and Landsford HL (1990), The regional acid deposition model and engineering model, NAPAP SOS/T Report 4, in National Acid Precipitation Assessment Program, Acidic Deposition: State of Science and Technology, Volume I, Washington, D.C.

[72] DeMore WB, Sander SP, Golden DM, Hampson RF, Kurylo MJ, Howard CJ, Ravishankara AR, Kolb CE, and Molina MJ (1994), Chemical kinetics and photochemical data for use in stratospheric modeling: Evaluation number 11, JPL Pub. 94-26, National Aeronautics and Space Administration, Jet Propulsion Laboratory, Pasadena, CA.

[73] Demerjian KL, Schere KL, and Peterson JT (1980), Theoretical estimates of actinic (spherically integrated) flux and photolytic rate constants of atmospheric species in the lower troposphere, in *Advances in Environmental Science and Technology*, Vol. 10, by John Wiley & Sons, Inc., 369-459.

[74] Elterman L (1968), UV, visible, and IR attenuation for altitudes to 50 km, AFCRL-68-0153, Air Force Cambridge Res. Lab. Bedford, MA.

[75] Stephens GL (1978), Radiation profiles in extended water clouds. II. Parameterization schemes, *J. Atmos. Sci.* **35**, 2123-2132.

[76] Gery MW, Whitten GZ, Killus JP, and Dodge MC (1989), A photochemical kinetics mechanism for urban and regional scale computer modeling, *J. Geophys. Res*. **94**, 12,925-12,956.

[77] Stockwell WR, Middleton P, and Chang JS (1990), The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.* **95**(D10), 16,343-16,367.

[78] Carter WPL (1994), Development of ozone reactivity scales for volatile organic compounds, *Journal of the Air and Waste Management Association*, 44, 881-899.

[79] Carter WPL (2000) Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment, Final report to the California Air Resources Board, Contract No. 92-329 and 95-308.

[80] Carter WPL and Atkinson R (1996), Development and evaluation of a detailed mechanism for the atmospheric reactions of isoprene and NOx, Int. J. Chem. Kinet. 28, 497-530.

[81] Carter WPL (1996), Condensed atmospheric photooxidation mechanisms for isoprene, *Atmos. Environ.* **24**, 4275-4290.

[82] Jacob DJ (2000), Heterogeneous chemistry and tropospheric ozone, *Atmos. Environ*., **34**, 2131-2159.

[83] Gong W and Cho HR (1993), A numerical scheme for the integration of the gas-phase chemical rate equations in three-dimensional atmospheric models, *Atmos. Environ.* **27A**, 2,147- 2,160.

[84] Gear CW (1971), *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ.

[85] Jacobson M and Turco RP (1994), SMVGEAR: A sparse-matrix, vectorized Gear code for atmospheric models, *Atmos. Environ.* **28**, 273-284.

[86] Carmichael GR, Peters LK, and Kitada T, A second generation model for regional-scale transport/chemistry/deposition, *Atmos. Environ.*, **20**, 173-188.

[87] Mathur R, Young JO, Schere KL, and Gipson GL (1998), A comparison of numerical techniques for solution of atmospheric kinetic equations, *Atmos. Environ.*, **32**, 1,535-1,553.

[88] Young JO, Sills E, and Jorge D (1993), Optimization of the Regional Oxidant Model for the Cray Y-MP, EPA/600/R-94-065, U.S. EPA, Research Triangle Park, NC.

[89] Hertel O, Berkowicz R, Christensen J, and Hov O (1993), Test of two numerical schemes for use in atmospheric transport-chemistry models, *Atmos. Environ.*, **27A**, 2591-2611.

[90] Huang HC and Chang JS (2001), On the performance of numerical solvers for a chemistry submodel in three-dimensional air quality models, Part 1 Box-model simulations, *J. Geophys. Res*., **188**, 20175-20188.

[91] Binkowski FS and Roselle SJ (2003), Models-3 Community Multiscale Air Quality (CMAQ] model aerosol component. I: Model description, *J. Geophys Res*., **108**(D6), 4183, doi:10.1029/2001JD001409.

[92] Binkowski FS and Shankar U (1995), The regional particulate model: I. Model description and preliminary results, *J. Geophys. Res.*, **100**(D12), 26191-26209.

[93] Nenes A, Pandis SN, Pilinis C (1998). ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols, *Aquat.Geoch*., **4**, 123-152.

[94] Nenes A., Pilinis C., and Pandis S.N. (1998) Continued Development and Testing of a New Thermodynamic Aerosol Module for Urban and Regional Air Quality Models, *Atmos. Env*., **33**, 1553-1560.

[95] Seigneur C, Huidischewski AB, Seinfeld JH, Whitby KT, Whitby ER, Brock JR, and Barnes HM (1986), Simulation of aerosol dynamics: A comparative review of mathematical models, *Aerosol Sci. and Technol*., **5**, 205-222.

[96] Whitby KT (1978), The physical characteristics of sulfur aerosols, *Atmos. Environ.*, **12**, 135-159.

[97] Dentener FJ and Crutzen PJ (1993), Reaction of N_2O_5 on the tropospheric aerosols: impact on the global distributions of NOx, O_3, and OH, *J. Geophys. Res*., **98**, 7149-7163.

[98] Riemer N, Vogel H, Vogel B, Schell B, Ackermann I, Kessler C and Hass H (2003), Reaction of N_2O_5 on the tropospheric aerosols: impact on the global distributions of NOx, O_3, and OH, *J. Geophys. Res*., **108**.

[99] Mentel TF, Sohn M, and Wahner A (1999), Nitrate effect on the heterogeneous hydrolysis of dintrogen pentoxide on aqueous aerosols, *J. Phys. Chem.*, **1**, 5451-5457.

[100] Schell B, Ackermann IJ, Hass H, Binkowski FS and Ebel, A (2001), Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, *J. Geophys. Res*., **106**, 28275-28293.

[101] Odum JR, Jungkamp TPW, Griffin RJ, Flagan RC and Seinfeld JH (1997), The atmospheric aerosol-forming potential of whole gasoline vapor. *Science*, **276**, 96-99.

[102] Griffin RJ, Cocker DR, Flagan RC and Seinfeld JH (1999), Organic aerosol formation from the oxidation of biogenic hydrocarbons, *J. Geophys. Res*., **104**, 3555-3567.

[103] Strader R, Lurmann F, and Pandis SN (1999), Evaluation of secondary organic aerosol formation in winter, *Atmos. Environ*., **33**, 4849-4863.

[104] Dennis RL, McHenry JN, Barchet WR, Binkowski FS, and Byun DW (1993), Correcting RADM's sulfate underprediction: Discovery and correction of model errors and testing the corrections through comparisons against field data, *Atmos. Environ.*, **26A**(6), 975-997.

[105] Walcek CJ and Taylor GR (1986), A theoretical method for computing vertical distributions of acidity and sulfate production within cumulus clouds, *J. Atmos. Sci* **43**, 339-355.

[106] Kim YP, Seinfeld JH, and Saxena P (1993), Atmospheric gas-aerosol equilibrium: I. Thermodynamics model, *Aerosol Sci. and Technol*., **19**, 157-181.

[107] Binkowski FS (1999), Aerosols in Models-3 CMAQ, Chapter 10, Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system, edited by Byun DW and Ching JKS, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[108] Gillani NV and Godowitch JM (1999), Plume-in-grid treatment of major point source emissions, Chapter 9, Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system, edited by Byun DW and Ching JKS, National Exposure Research Laboratory, U.S. EPA, Research Triangle Park, NC.

[109] Godowitch JM and Young JO (2000), Photochemical simulations of point source emissions with the Models-3 CMAQ plume-in-grid approach, A&WMA 91st Annual Meeting, Salt Lake City, UT.

[110] Godowitch JM (2001), Results of photochemical simulations of subgrid scale point source emissions with the Models-3 CMAQ modeling system, in the Millenium Symposium on Atmospheric Chemistry, Proceedings of the American Meteorological Society, January 14-18, Albuquerque, NM, 43-49.

[111] U.S. EPA (2003), 1999 National Emission Inventory Documentation and Data, available at http://www.epa.gov/ttn/chief/net/1999inventory.html

[112] TexAQS 200 web site: http://www.utexas.edu/research/ceer/texaqs/participants/about.html

[113] Allen DT, Estes M, Smith J and Jeffries H (2002), Accelerated Science Evaluation of Ozone Formation in the Houston-Galveston Area: Overview, available at http://www.utexas.edu/research/ceer/texaqsarchive

[114] Nielsen-Gammon JW (2002), Meteorological Modeling for the August 2000 Houston-Galveston Ozone Episode: METSTAT Statistical Evaluation and Model Runs from March-June 2002. Report to the Technical Analysis Division, Texas Natural Resource Conservation Commission. June 2002.

Table 1. Vertical coordinates in atmospheric models and their associated characteristics.

Table 2. List of science process classes and subroutines called by the CMAQ driver.

Table 3. List of species considered in the CMAQ aqueous chemistry module.

Table 4. Model evaluation statistics for daily maximum 8-hr average ozone concentrations (ppb) over the period June 15-July 16, 1999. Table indicates mean, standard deviation (SD), coefficient of variation (CV), and percentile levels in the CMAQ model and AIRS observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME).

Table 5. Model evaluation statistics for daily average PM2.5 concentrations ($\mu g\ m^{-3}$) over the period June 15-July 16, 1999. Table indicates mean and percentile levels in the CMAQ model and IMPROVE network observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME).

Table 6. Model evaluation statistics for daily average PM2.5 concentrations ($\mu g\ m^{-3}$) over the period January 4-February 19, 2002. Table indicates mean and percentile levels in the CMAQ model for IMPROVE and STN network observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME) for IMPROVE and STN comparisons.

Table 1. Vertical coordinates in atmospheric models and their associated characteristics

CMAQ Coordinate $\hat{x}^3 = \xi$	Meteorological Coordinate Definition	Contravariant Vertical Velocity $\hat{v}^3$	Vertical Jacobian J_ξ	Remarks
z		$w = \frac{dz}{dt}$	1	geometric height
$\sigma_{\tilde{z}}$	$\sigma_{\tilde{z}} = H\frac{z - z_{sfc}}{H - z_{sfc}}$	$\frac{d\sigma_{\tilde{z}}}{dt}$	$\frac{H - z_{sfc}}{H}$	H is the thickness of model and $\sigma_{\tilde{z}}$ is the scaled height
σ_z	$\sigma_z = \frac{z - z_{sfc}}{H - z_{sfc}}$	$\frac{d\sigma_z}{dt}$	$H - z_{sfc}$	nondimensional height, terrain-influenced
$1 - \sigma_{po}$	$\sigma_{po} = \frac{p_o - p_T}{p_{os} - p_T}$	$-\frac{d\sigma_{po}}{dt}$	$\frac{p_{os} - p_T}{\rho_o g}$	nondimensional reference pressure
$1 - \sigma_{\tilde{p}}$	$\sigma_{\tilde{p}} = \frac{\tilde{p} - p_T}{\tilde{p}_s - p_T}$	$-\frac{d\sigma_{\tilde{p}}}{dt}$	$\frac{\tilde{p}_s - p_T}{\tilde{\rho} g}$	nondimensional hydrostatic pressure
$1 - \eta$	$\eta = \left(\frac{\tilde{p} - p_T}{\tilde{p}_s - p_T}\right)\eta_{sfc}$	$-\frac{d\eta}{dt}$	$\frac{\tilde{p}_s - \tilde{p}_T}{\tilde{\rho} g \eta_{sfc}}$	step-mountain ETA $\eta_{sfc} = \left(\frac{p_o(z_{sfc}) - p_T}{p_o(0) - p_T}\right)$

z_{sfc} : height of topography from mean sea level

H : thickness of model

p_o : reference pressure at height z

p_{os} : reference pressure at z_{sfc}

p_T : pressure at model top

$\tilde{p}$: hydrostatic pressure at height z

$\tilde{p}_s$: hydrostatic pressure at z_{sfc}

Table 2. List of science process subroutines called by the CMAQ Driver

Subroutine called by DRIVER[&]	Science Class	Description
CGRID_MAP	UTIL	Sets up pointers for different concentration species: gas chemistry, aerosol, non-reactive, and tracer species
INITSCEN	INIT[*]	Initializes simulation time period, time stepping constants, and concentration arrays for the driver
ADVSTEP	DRIVER	Computes the model synchronization time step and number of repetitions for the output time step
COUPLE/ DECOUPLE	COUPLE[*]	Converts units and couples or de-couples concentration values with the density and Jacobian for transport
SCIPROC	DRIVER	Controls all of the physical and chemical processes for a grid (currently, two versions are available: symmetric and asymmetric around the chemistry processes)
XADV, YADV	HADV	Computes advection in horizontal plane (x- and y-directions)
ZADV	VADV	Computes advection in the vertical direction in the generalized coordinate system
ADJADV	ADJCON	Adjusts concentration fields to ensure mixing ratio conservation given mass consistency error in meteorology data
HDIFF	HDIFF	Computes horizontal diffusion
VDIFF	VDIFF	Computes vertical diffusion and deposition
CHEM	CHEM	Solves gas-phase chemistry
PING	PING	Computes effects of plume-in-grid process
AERO	AERO	Computes aerosol dynamics, particle formation, and deposition
CLDPRC	CLOUD	Computes cloud mixing and aqueous chemistry
PA_UPDATE	PROCAN	Computes amount of concentration change of each process call

[&]CMAQ includes other modules that are not called by driver. Thye are mostly classes for modules to compute necessary interim paramters, such as PHOT for modification of photolysis rates for cloud attenuation and AERO_DEPV for estimating particl size dependent dry deposition velocites used by AERO.

[*]Represents a process class that is part of DRIVER function.

Table 3. List of species considered in the CMAQ aqueous chemistry module

Gases:	**Aerosols:**
SO_2	$SO_4^{=}$ (Aitken & accumulation modes)
HNO_3	NH_4^{+} (Aitken & accumulation modes)
N_2O_5	NO_3^{-} (Aitken, accumulation, & coarse modes)
CO_2	Organics (Aitken & accumulation modes)
NH_3	Primary (Aitken, accumulation, & coarse modes)
H_2O_2	$CaCO_3$
O_3	$MgCO_3$
HCOOH	NaCl
$CH_3(CO)OOH$	Fe^{3+}
CH_3OOH	Mn^{2+}
H_2SO_4	KCl
	Number (Aitken, accumulation, & coarse modes)
	Surface Area (Aitken & accumulation modes)

Table 4. Model evaluation statistics for daily maximum 8-hr average ozone concentrations (ppb) over the period June 15-July 16, 1999. Table indicates mean, standard deviation (SD), coefficient of variation (CV), and percentile levels in the CMAQ model and AIRS observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME).

		AIRS		
	CMAQ	**OBS**		
Mean	54.5	50.2	n	23,196
SD	13.7	18.2	R	0.75
CV	25.1%	36.3%	MB	4.3
95th	77.2	82.0	NMB(%)	8.7
50th	54.1	48.9	RMSE	12.8
5th	33.4	22.5	NME(%)	20.2

Table 5. Model evaluation statistics for daily average PM2.5 concentrations (μg m^{-3}) over the period June 15-July 16, 1999. Table indicates mean and percentile levels in the CMAQ model and IMPROVE network observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME).

	CMAQ	**IMPROVE**		
			n	457
Mean	6.75	7.48	R	0.71
			MB	-0.73
95th	20.2	19.10	NMB(%)	-9.8
50th	4.49	5.22	RMSE	4.70
5th	1.36	1.78	NME(%)	40.3

Table 6. Model evaluation statistics for daily average PM2.5 concentrations (μg m^{-3}) over the period January 4-February 19, 2002. Table indicates mean and percentile levels in the CMAQ model for IMPROVE and STN network observations. Also indicated are the number of data points (n), correlation coefficient (R), mean and normalized mean bias (MB, NMB), and root mean square and normalized mean error (RMSE, NME) for IMPROVE and STN comparisons.

	CMAQ	**IMPROVE**	**CMAQ**	**STN**	**IMPROVE**		**STN**	
					n	714	n	927
Mean	5.20	3.71	12.96	12.45	R	0.68	R	0.37
					MB	1.49	MB	0.51
95th	14.8	11.7	26.9	27.2	NMB(%)	40.3	NMB(%)	4.1
50th	3.37	2.25	11.0	10.0	RMSE	3.81	RMSE	10.45
5th	0.95	0.53	3.35	3.8	NME(%)	68.9	NME(%)	50.0

Figure 1. Science process modules in Models-3 CMAQ. Independent processors are represented with round rectangles and interface processes are shown with rectangular boxes. Typical science process modules (in hexagon boxes) update the concentration field directly and the data-provider modules (e.g., Photolysis routine in a pentagon box) include routines to feed appropriate environmental input data to the science process modules.

Figure 2. Driver module and its science process call sequence.

Figure 3. Grid-size dependent horizontal diffusivity used in CMAQ (estimated for different magnitude of deformation, from 10^{-6} to 10^{-3} s^{-1}).

Figure 4. Evolution of aerosol size distributions for the clear, urban, and hazy cases. Initial conditions of *Seigneur et al* [95] are used.

Figure 5. Comparison of PinG modeled species concentrations for ozone (solid line), NO_y (thick solid line) and SO_2 (long dash) versus observed plume concentrations obtained from a horizontal aircraft traverse intercepting multiple point source plumes from Johnsonville (extreme left plume) and the Cumberland plume (in the middle) near Nashville, Tennessee, 18:45 UTC, 7 July, 1995.

Figure 6. Scatterplot of daily maximum 8-hr average ozone concentrations (ppb) from continental U.S. AIRS stations for the period June 15-July 16, 1999 versus comparable CMAQ model estimates. Solid lines are 2:1 and 1:2 boundaries and dotted line is best fit to data.

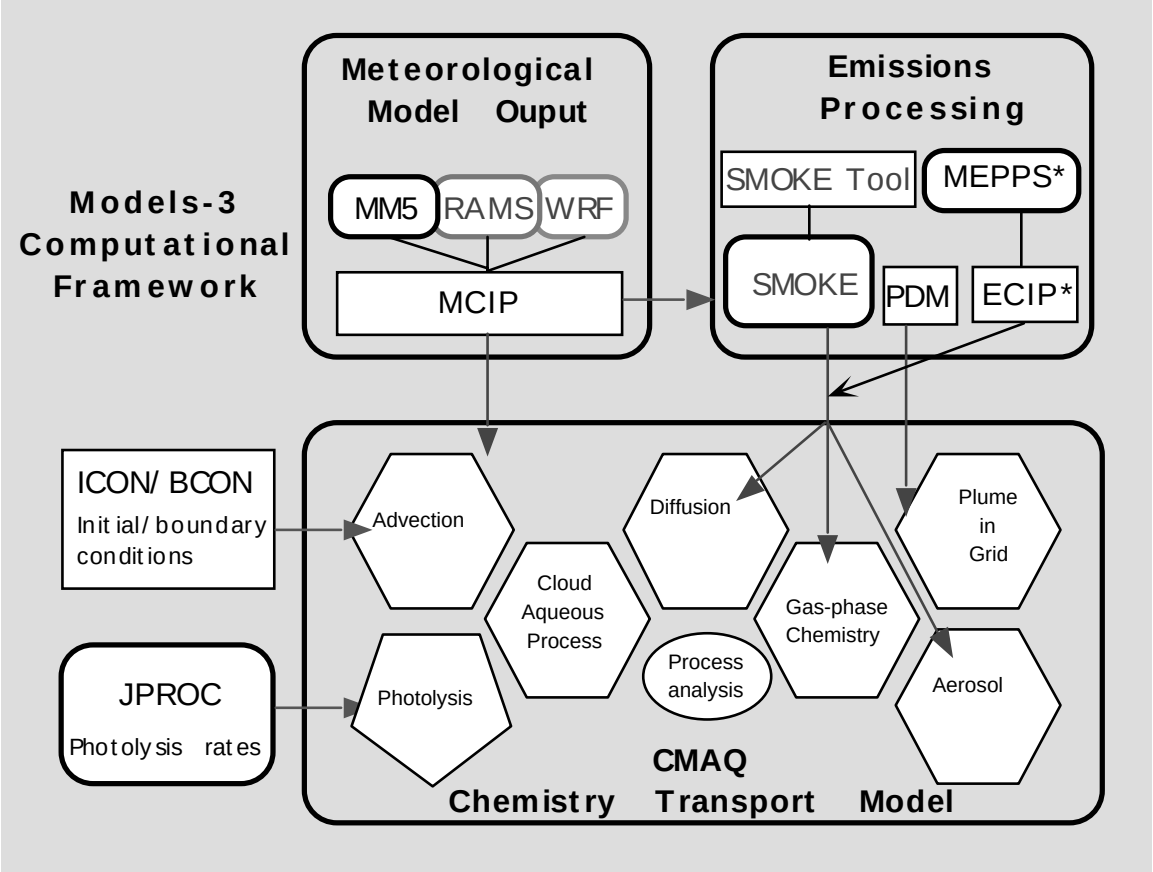
Figure 7. Scatterplot of daily average PM2.5 concentrations ($\mu g\ m^{-3}$) from continental U.S. monitoring stations for the period June 15-July 16, 1999 versus comparable CMAQ model estimates. Results are color-coded according to eastern or western U.S. Solid lines are 1:1 as well as 2:1 and 1:2.

Figure 8. Scatterplot of daily average PM2.5 concentrations ($\mu g\ m^{-3}$) from continental U.S. monitoring stations for the period January 4-February 19, 2002 versus comparable CMAQ model estimates. Results are color-coded according to eastern or western U.S. Solid lines are 1:1 as well as 2:1 and 1:2.

Figure 9. Daily maximum modeled ozone concentrations (ppm) with modified VOC emissions data.

Figure 10. Scatter diagram comparing CAMS and modeled ozone concentrations (ppb) with modified VOC emissions data.

Figure 11. Time series of CAMS and modeled ozone concentrations (ppb) with modified VOC emissions data for non-industrial sites.



* Used in versions of CMAQ released before 2001

Figure 1

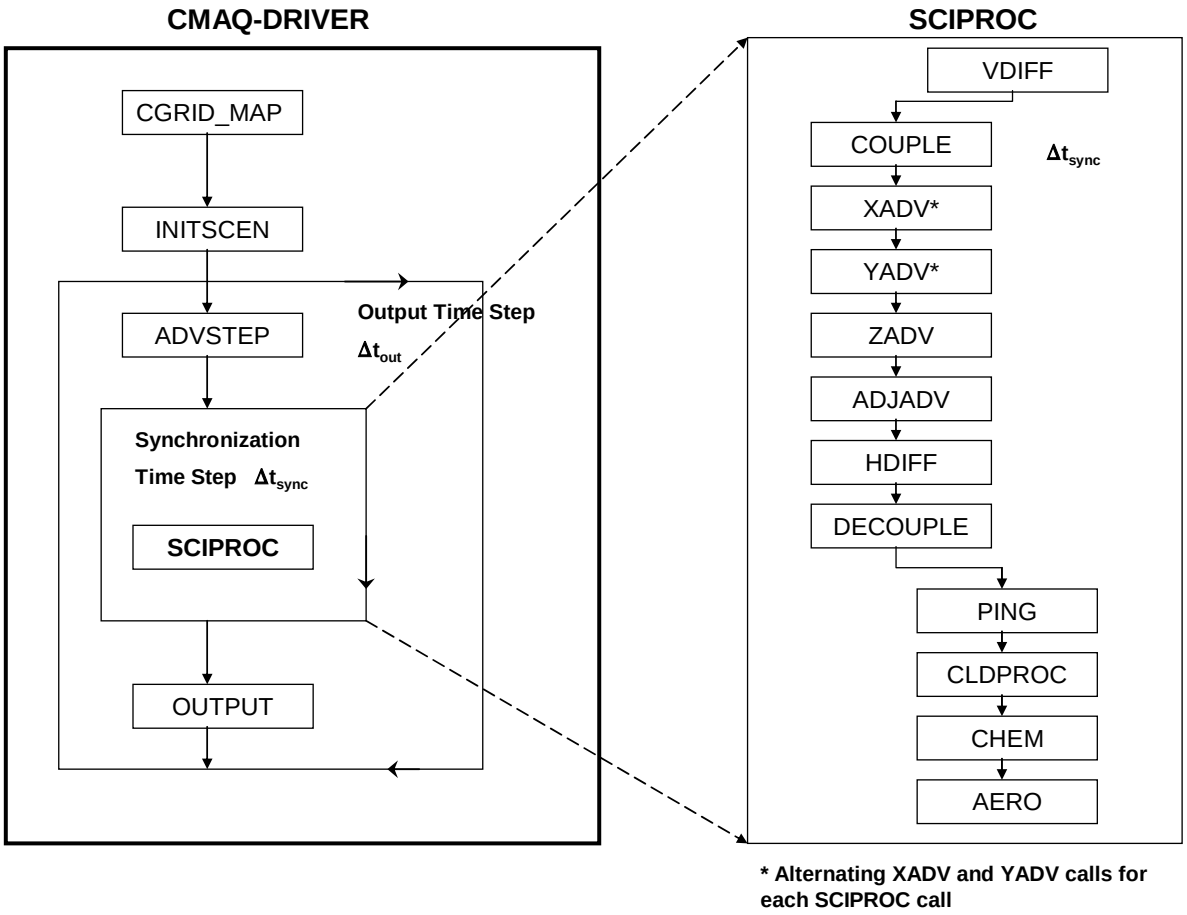

CMAQ-DRIVER
CGRID_MAP
INITSCEN
ADVSTEP
Output Time Step
Δt_out
Synchronization
Time Step Δt_sync
SCIPROC
OUTPUT
SCIPROC
VDIFF
COUPLE
Δt_sync
XADV*
YADV*
ZADV
ADJADV
HDIFF
DECOUPLE
PING
CLDPROC
CHEM
AERO
* Alternating XADV and YADV calls for each SCIPROC call

Figure 2

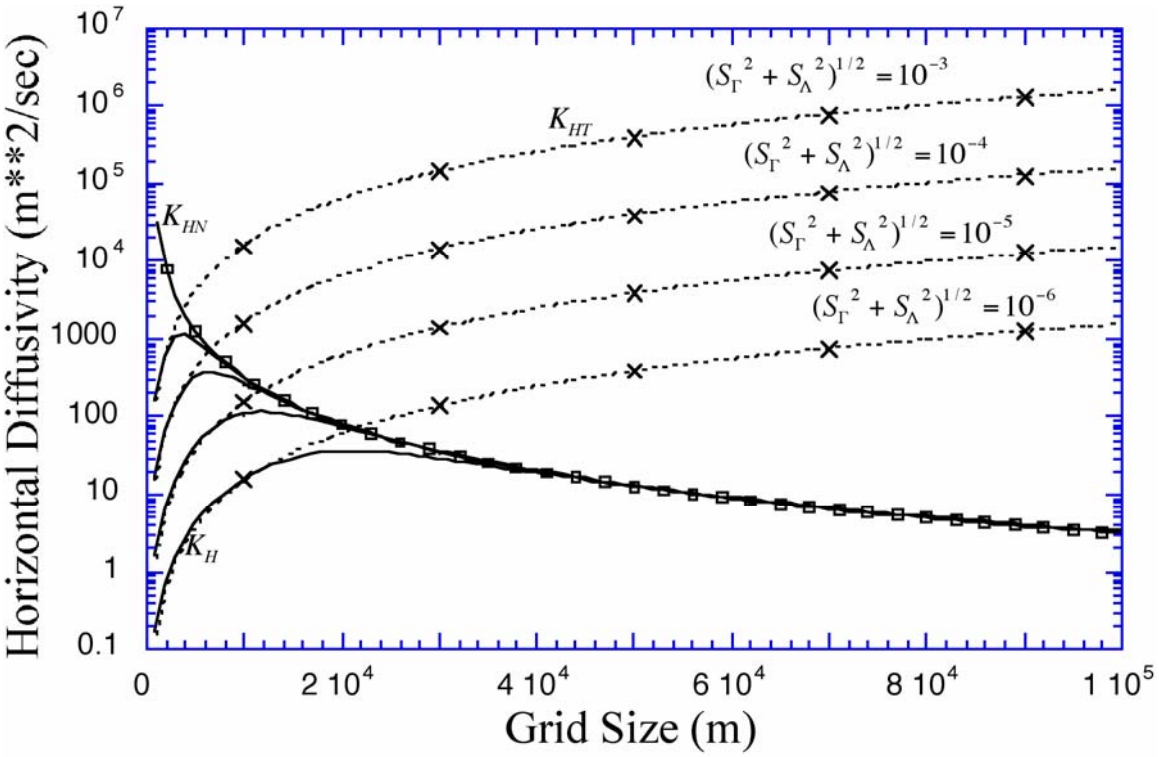

Horizontal Diffusivity (m**2/sec)
Grid Size (m)
10^7
10^6
10^5
10^4
1000
100
10
1
0.1
0
$2\ 10^4$
$4\ 10^4$
$6\ 10^4$
$8\ 10^4$
$1\ 10^5$
K_{HT}
K_{HN}
K_H
$(S_\Gamma^2 + S_\Lambda^2)^{1/2} = 10^{-3}$
$(S_\Gamma^2 + S_\Lambda^2)^{1/2} = 10^{-4}$
$(S_\Gamma^2 + S_\Lambda^2)^{1/2} = 10^{-5}$
$(S_\Gamma^2 + S_\Lambda^2)^{1/2} = 10^{-6}$

Figure 3

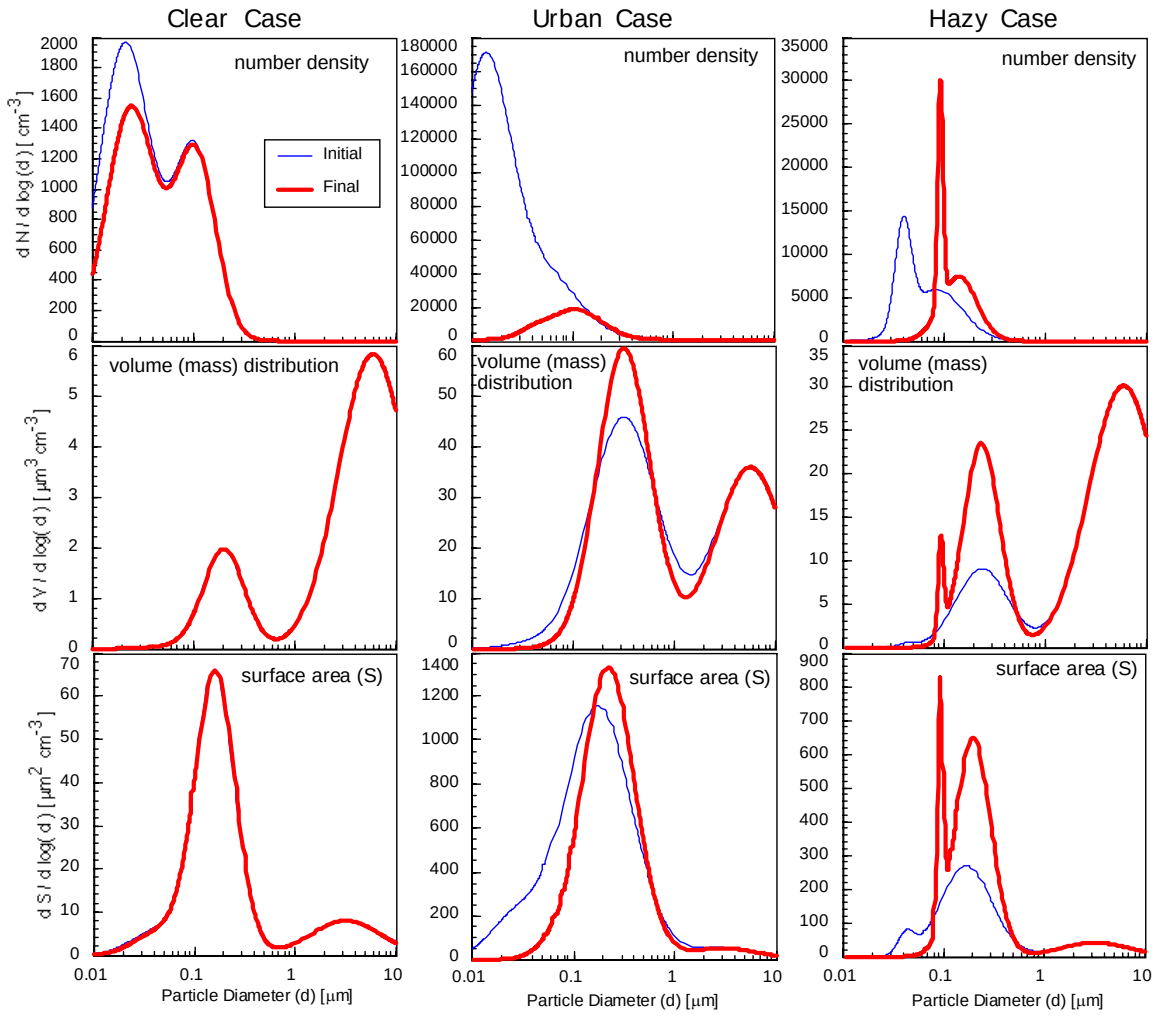

Clear Case
Urban Case
Hazy Case
number density
Initial
Final
volume (mass) distribution
surface area (S)
d N / d log (d) [cm-3]
d V / d log(d) [µm3 cm-3]
d S / d log(d) [µm2 cm-3]
Particle Diameter (d) [µm]

Figure 4

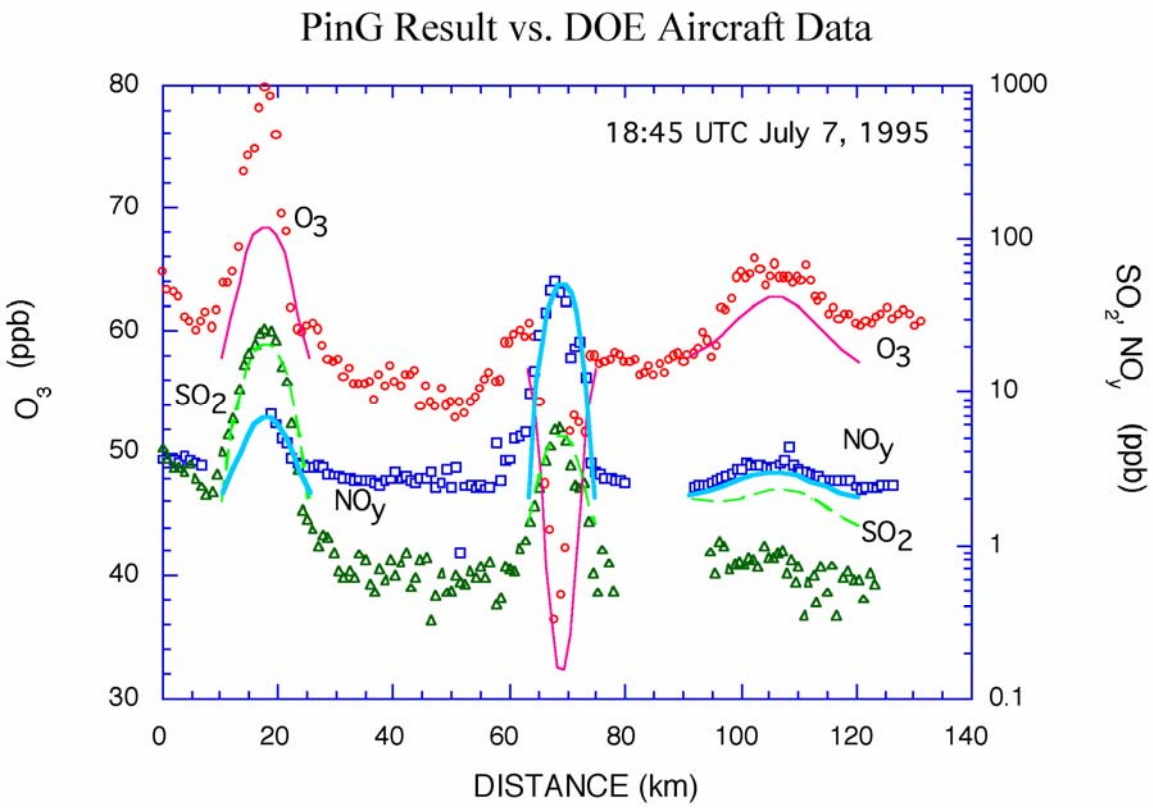

PinG Result vs. DOE Aircraft Data
18:45 UTC July 7, 1995
O3
SO2
NOy
O3
NOy
SO2
O3 (ppb)
SO2, NOy (ppb)
DISTANCE (km)
80
70
60
50
40
30
1000
100
10
1
0.1
0
20
40
60
80
100
120
140

Figure 5

Figure 6

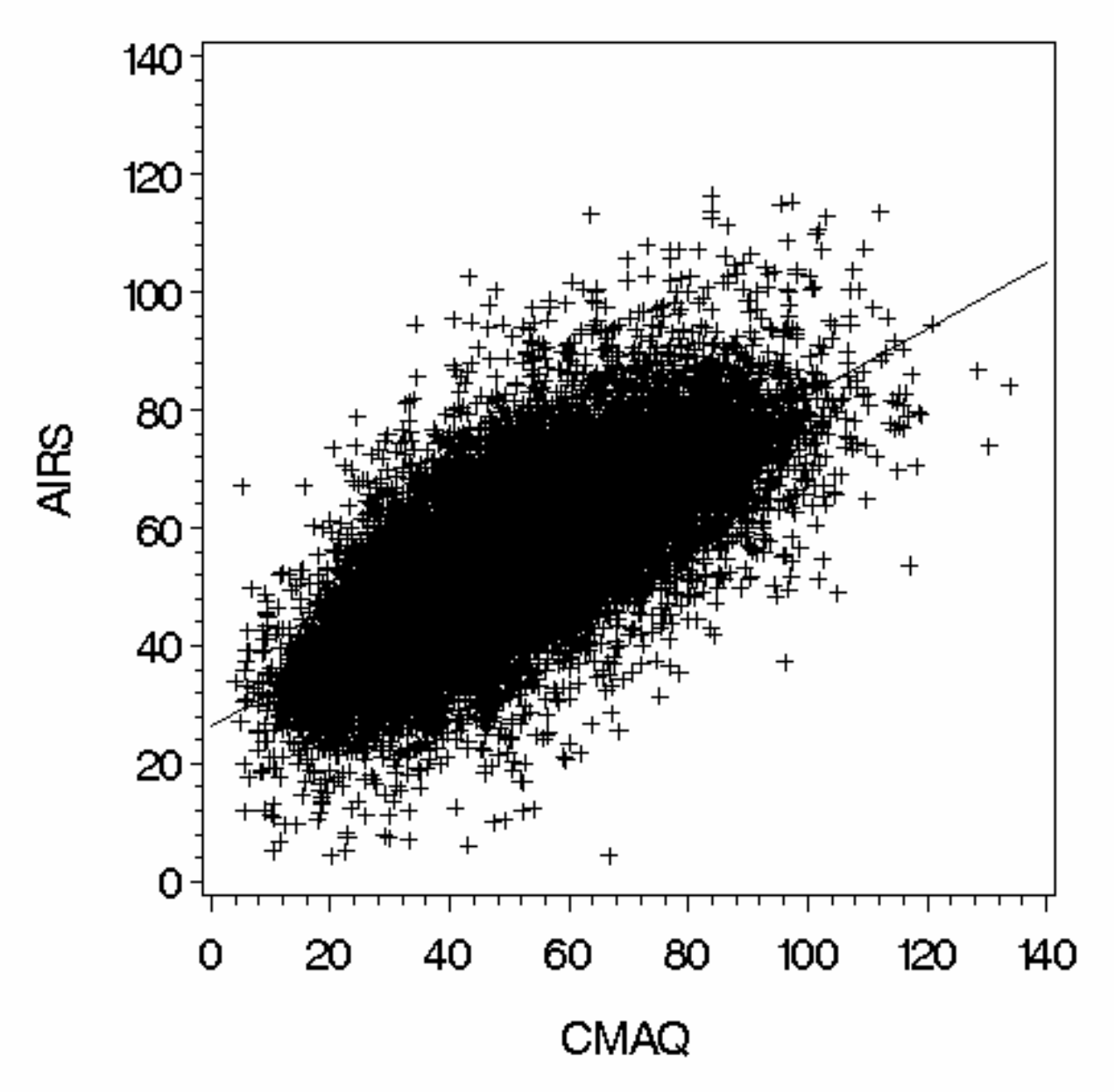


Figure 7

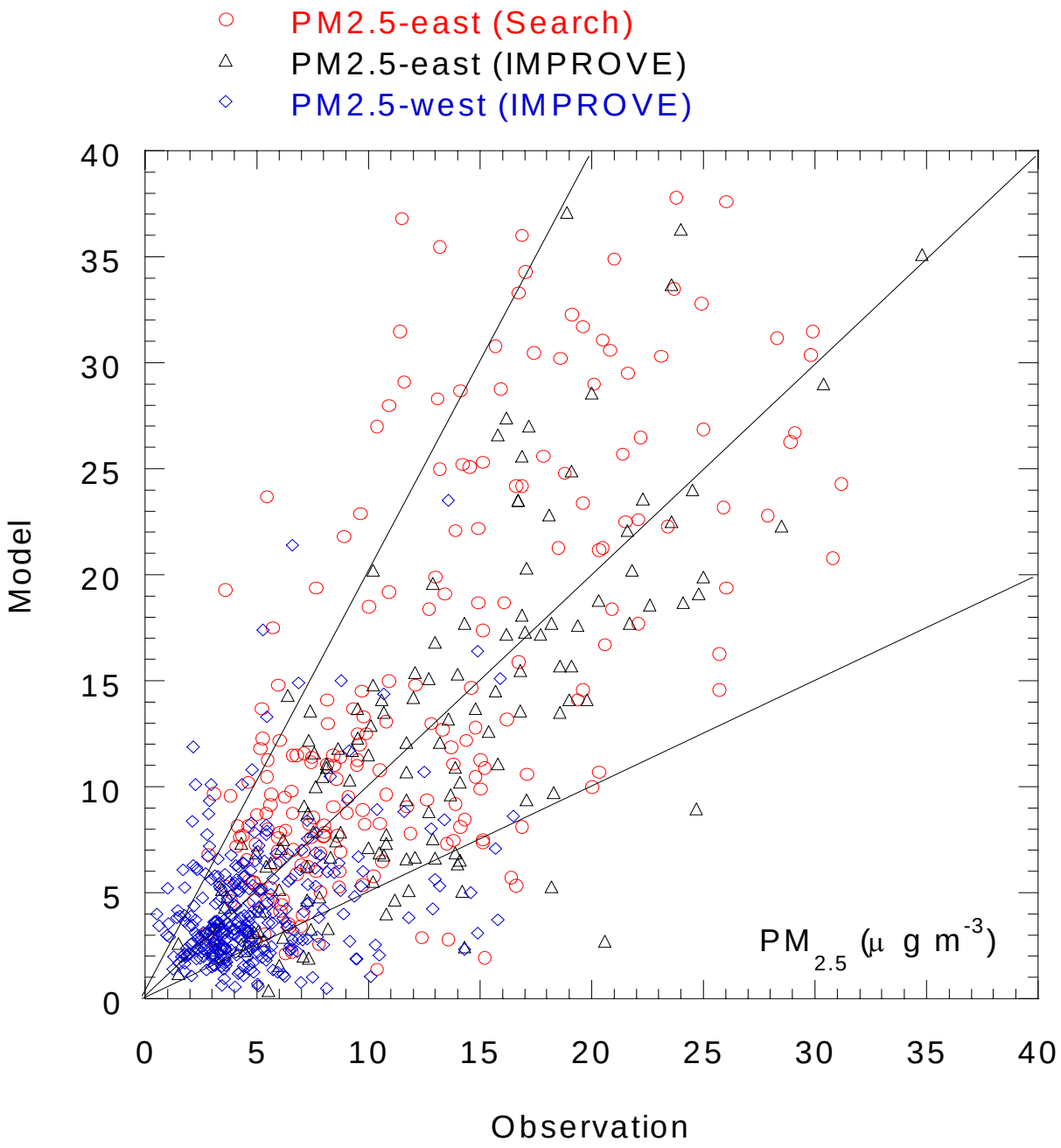

PM2.5-east (Search)
PM2.5-east (IMPROVE)
PM2.5-west (IMPROVE)
Model
Observation
0
5
10
15
20
25
30
35
40
PM2.5 (μ g m-3)

Figure 8

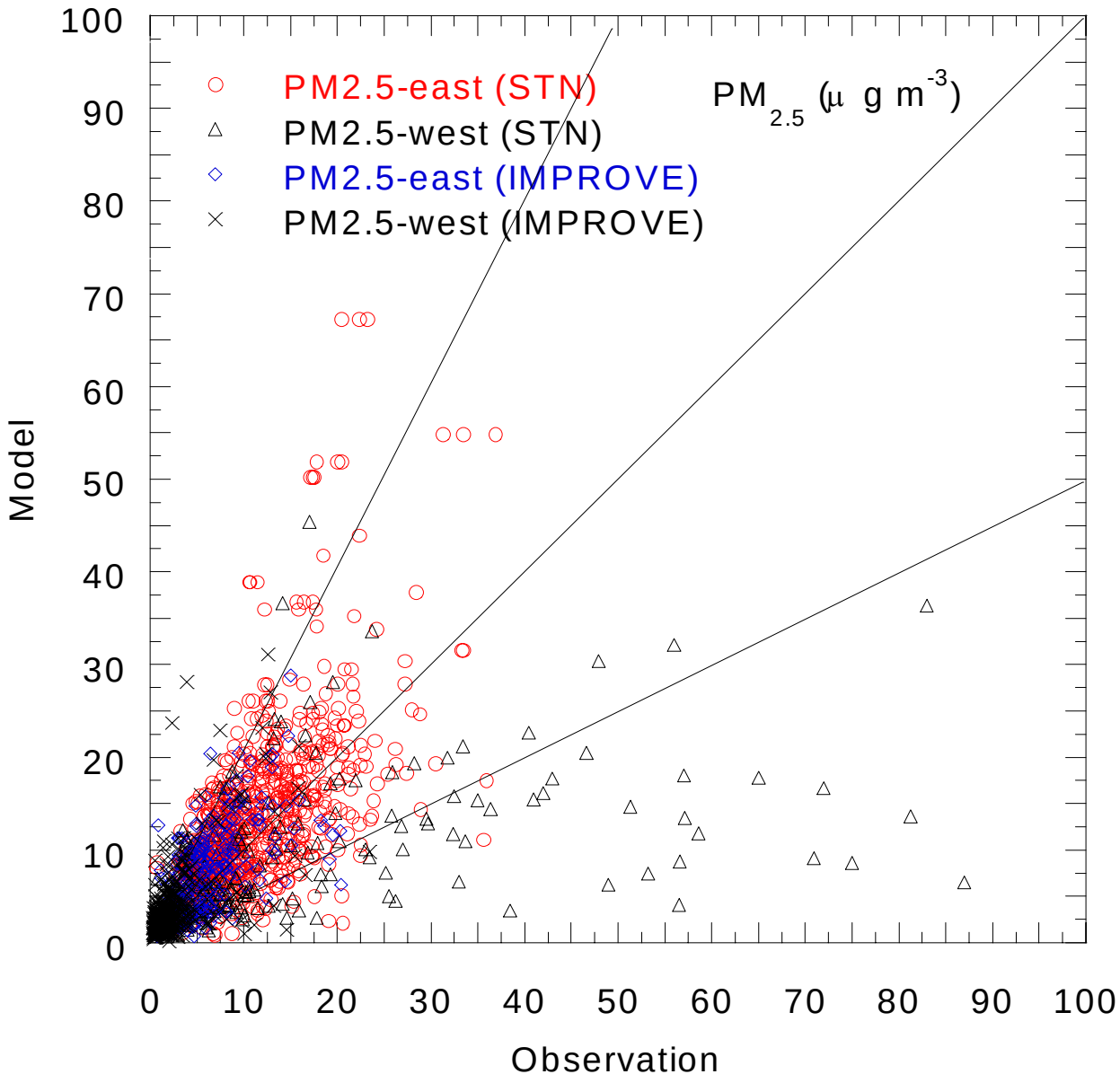

PM2.5-east (STN)
PM2.5-west (STN)
PM2.5-east (IMPROVE)
PM2.5-west (IMPROVE)
PM2.5 (μ g m-3)
Model
Observation
0
10
20
30
40
50
60
70
80
90
100

Figure 9

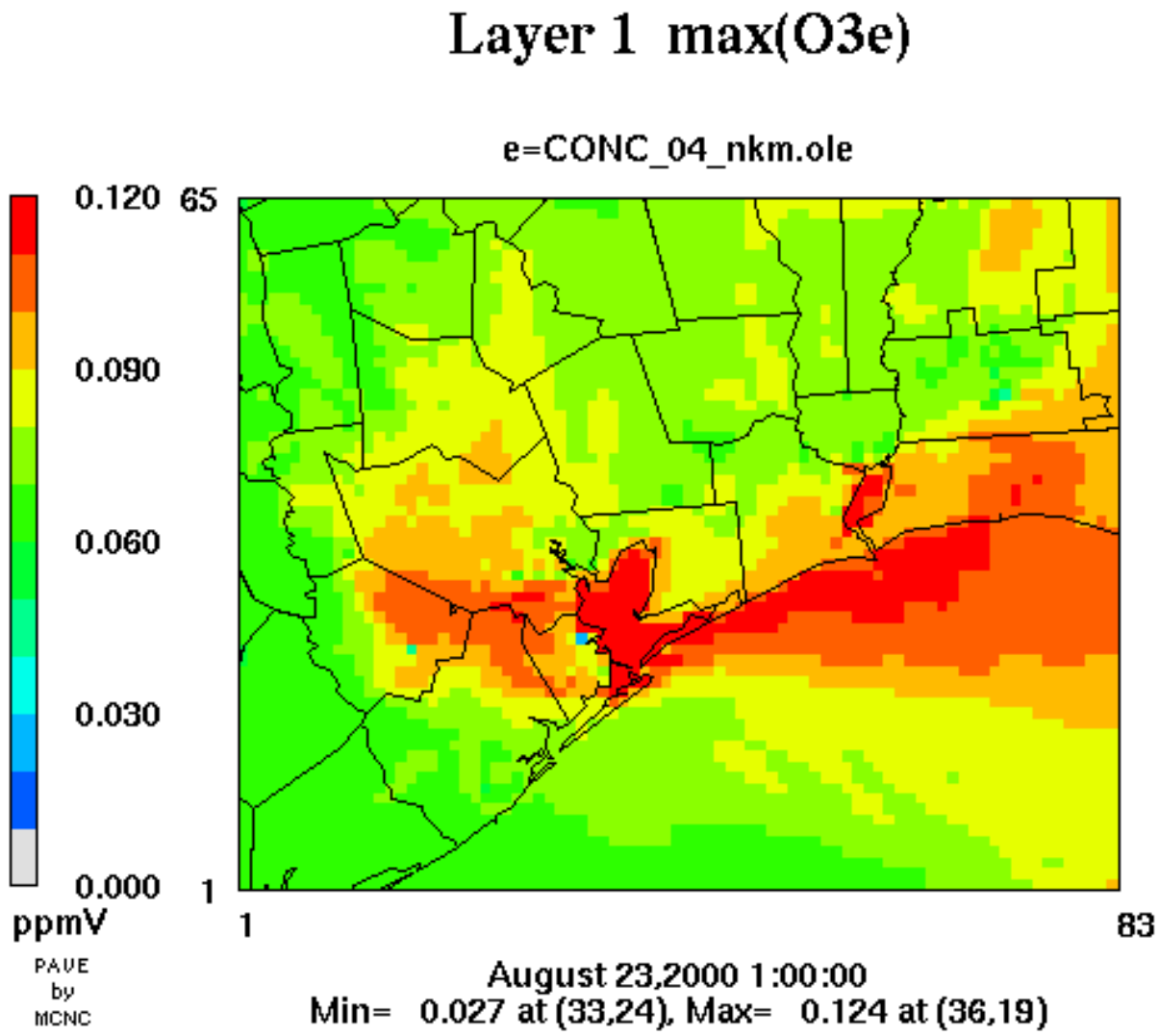

Layer 1 max(O3e)
e=CONC_04_nkm.ole
0.120
0.090
0.060
0.030
0.000
ppmV
65
1
1
83
PAVE
by
MCNC
August 23,2000 1:00:00
Min= 0.027 at (33,24), Max= 0.124 at (36,19)

Figure 10

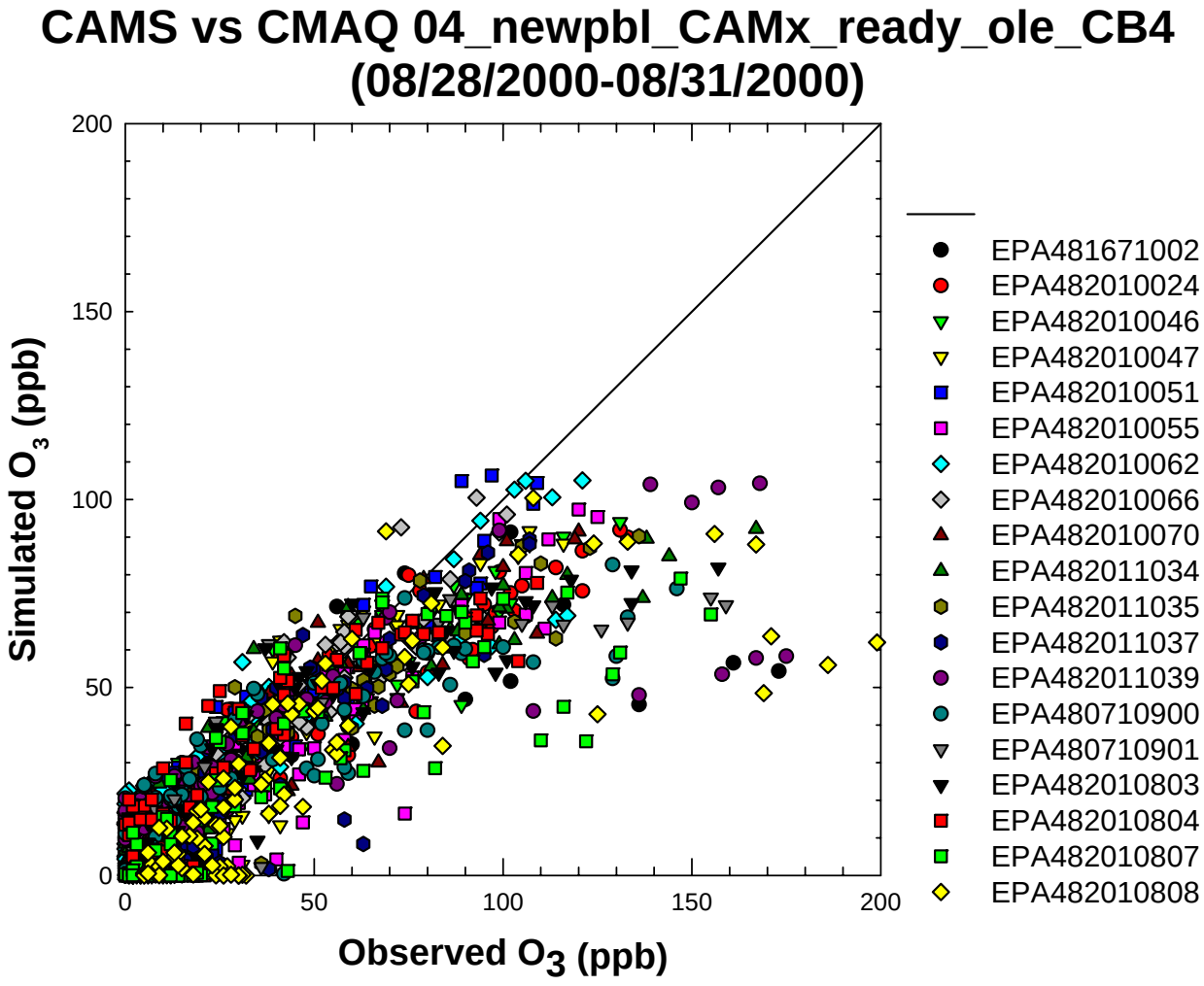
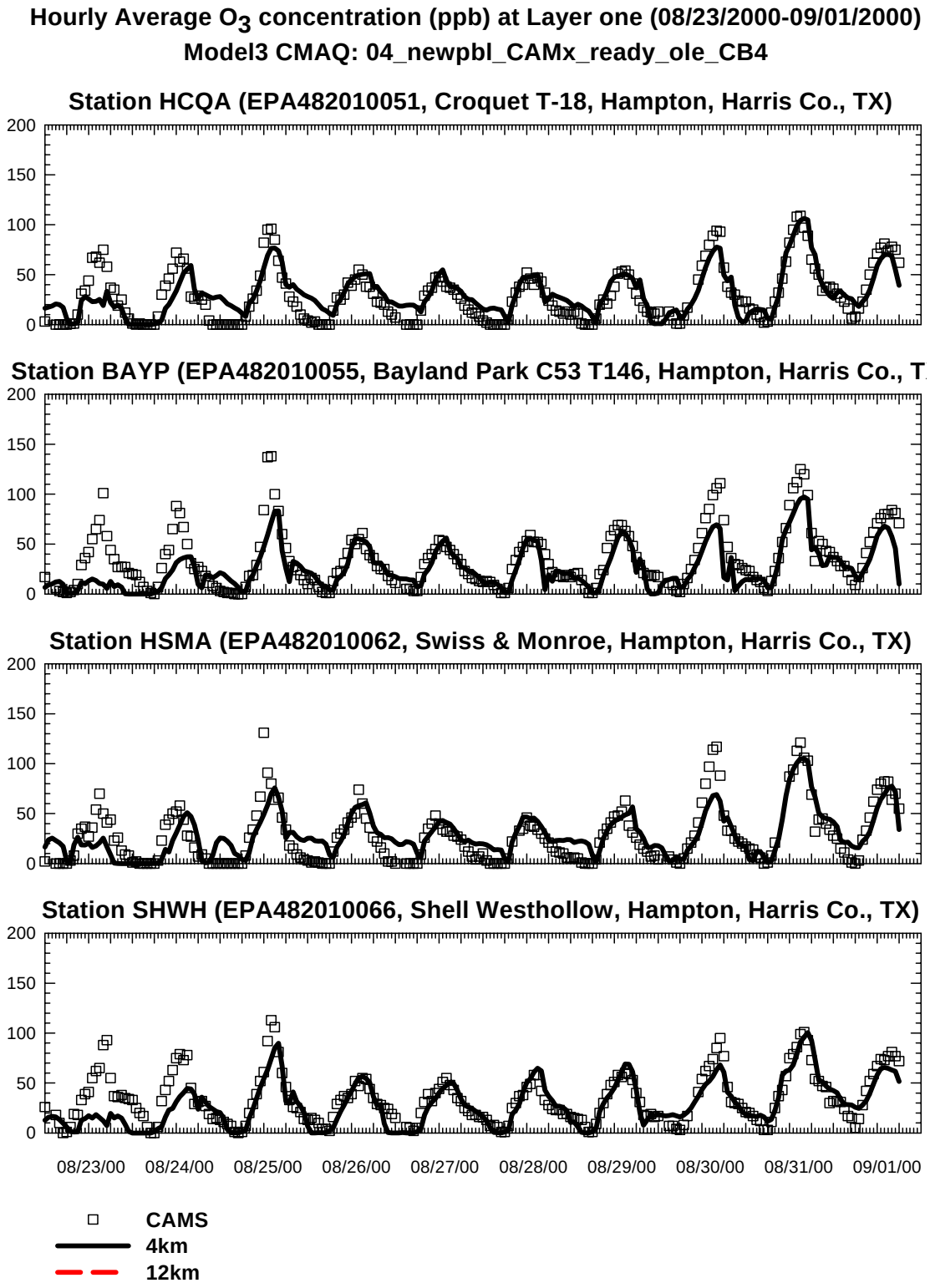

CAMS vs CMAQ 04_newpbl_CAMx_ready_ole_CB4
(08/28/2000-08/31/2000)
Simulated O3 (ppb)
Observed O3 (ppb)
0
50
100
150
200
EPA481671002
EPA482010024
EPA482010046
EPA482010047
EPA482010051
EPA482010055
EPA482010062
EPA482010066
EPA482010070
EPA482011034
EPA482011035
EPA482011037
EPA482011039
EPA480710900
EPA480710901
EPA482010803
EPA482010804
EPA482010807
EPA482010808

Figure 11


Hourly Average O_3 concentration (ppb) at Layer one (08/23/2000-09/01/2000)
Model3 CMAQ: 04_newpbl_CAMx_ready_ole_CB4
Station HCQA (EPA482010051, Croquet T-18, Hampton, Harris Co., TX)
Station BAYP (EPA482010055, Bayland Park C53 T146, Hampton, Harris Co., T
Station HSMA (EPA482010062, Swiss & Monroe, Hampton, Harris Co., TX)
Station SHWH (EPA482010066, Shell Westhollow, Hampton, Harris Co., TX)
200
150
100
50
0
08/23/00
08/24/00
08/25/00
08/26/00
08/27/00
08/28/00
08/29/00
08/30/00
08/31/00
09/01/00
CAMS
4km
12km